MULTIPOINT SCHUR ALGORITHM AND ORTHOGONAL RATIONAL FUNCTIONS: CONVERGENCE PROPERTIES.

L. BARATCHART, S. KUPIN, V. LUNOT, M. OLIVI

Abstract. Classical Schur analysis is intimately connected to the theory of orthogonal polynomials on the circle [43]. We investigate here the connection between multipoint Schur analysis and orthogonal rational functions. Specifically, we study the convergence of the Wall rational functions *via* the development of a rational analogue to the Szegő theory, in the case where the interpolation points may accumulate on the unit circle. This leads us to generalize results from [22, 10], and yields asymptotics of a novel type.

Introduction

The theory of orthogonal polynomials, with respect to a positive measure on the line or the circle, currently undergoes a period of intensive growth. To hint at recent advances, let us quote the papers by Killip-Simon [24], Martínez-Finkelstein et al. [32], Miña-Dìaz [34], Kuijlaars et al. [27], McLaughlin-Miller, [31], Lubinsky [29] and Remling [41]. A comprehensive account of many late developments in the field can be found in the monograph by Simon [43]. Let us mention in passing that, over the same period, non-Hermitian orthogonality with respect to complex measures, which is intimately connected with rational approximation and interpolation, made some progress too; see, for example, Aptekarev [4], Aptekarev-Van Assche [5], Baratchart-Küstner-Totik [7] and Baratchart-Yattselev [8].

The connection between orthogonal polynomials on the unit circle and the Schur algorithm is an old one. Recall that a Schur function is an analytic map from the open unit disk into itself. The Schur algorithm, introduced by Schur and Nevanlinna [42, 35], associates to every Schur function a sequence of complex numbers of modulus at most one, called its Schur (or Verblunsky) parameters. Since the *mise en scène* of the present article unfolds mainly in the framework of the Schur analysis, we shall call the parameters "Schur", although the term "Verblunsky" seems to be more fair historically, see Simon [43, Sect. 1.1] for a discussion. These parameters may be viewed as hyperbolic analogues of the Taylor coefficients at the origin. They generate a continued-fraction expansion of the function, whose truncations give rise to the so-called Schur approximants. These are hyperbolic counterparts of the

Date: January, 25, 2010.

1991 *Mathematics Subject Classification.* Primary: 30B70. Secondary: 41A20.

Key words and phrases. Approximation by rational functions, Schur algorithm, Schur (Verblunsky) parameters, orthogonal rational functions (orthogonal polynomials), Wall orthogonal functions (Wall polynomials).

This work was partially supported by grants ANR-07-BLAN-024701 and ANR-09-BLAN-005801.

Taylor polynomials, see definition (0.2) to come. Now, an elementary linear fractional transformation puts Schur functions in one-to-one correspondence with Carathéodory functions, *i.e.* analytic functions with positive real part in the disk, which are themselves in bijection with positive measures on the circle *via* the Herglotz transform. A long time ago already, Geronimus and Wall observed the remarkable identity between the Schur parameters of a function and the recurrence coefficients of the orthogonal polynomials associated to the corresponding measure [19, 45]. However, only relatively recently was it stressed by Khrushchev [22, 23] how properties of the measure, that govern the convergence of the corresponding orthogonal polynomials, are linked to the convergence of the Schur approximants *on* the unit circle.

It must be pointed out that the Schur algorithm is among the seldom procedures preserving the Schur character in rational approximation; equivalently, it yields Carathéodory rational approximants to Carathéodory functions on the disk or the half-plane. This feature is of fundamental importance in several areas of Physics and Engineering, where the Schur or Carathéodory nature of a transfer function is to be interpreted as a passivity property of the underlying system. Moreover, in such modeling issues, the relevant norms take place on the boundary of the analyticity domain, that is, on the circle or the line, see *e.g.* [33, 17, 3, 10]. This is why the results by Khrushchev are of significance from the applied viewpoint as well, which was one incentive for the authors to undertake the present study. This motivation is illustrated in the doctoral work by V. Lunot [30].

Unless the Schur function to be approximated possesses some symmetry, though, there is no particular reason why Schur approximants should distinguish the origin. It is thus natural to turn to multipoint Schur approximants, that play the role of Lagrange interpolating polynomials in the present hyperbolic context, see definitions (0.1) and (0.2) to come. The role of orthogonal polynomials is then played by orthogonal *rational* functions with poles at the reflections of the interpolation points across the unit circle. Orthogonal rational functions, pioneered by Dzrbasjan [13], were later studied by Pan [38] and considerably expanded by Bultheel et al. [10], see also Langer-Lasarow [28]. The last two references stress the connection with the multipoint Schur algorithm, and the comprehensive exposition in [10], which contains further references, presents an account of Szegő asymptotics when the interpolation points are compactly supported in the disk.

The present article is concerned with the so-called determinate case (see condition (0.3)) when the interpolation sequence may have limit points on the circle, and its purpose is two-fold. On the one hand, we derive analogues of Khrushchev's results [22] on the convergence of Schur approximants in the multipoint case, and on the other hand we present a counterpart of the Szegő theory for the associated orthogonal rational functions. We limit ourselves to regular measures on the circle, whose density does not vanish at limit points of the interpolation sequence, and we do not touch upon what is perhaps the most important issue, namely how to choose the interpolation points in an optimal fashion as regards convergence rates. Nonetheless, the present paper seems first to propose asymptotics when the interpolation points approach the unit circle.

Anyone writing on the subject faces the difficulty of expounding the maze of formulas on which research can dwell. Our choice has been to give a terse summary of what we use, along with references.

0.1. **Definitions.** Let $\mathbb{D}$ be the open unit disk and $\mathbb{T}$ the unit circle. A function f is called *Schur* if it belongs to the unit ball of the Hardy space $H^\infty(\mathbb{D})$, *i.e.* if $f \in H^\infty(\mathbb{D})$ and $||f||_\infty \leq 1$. The collection of Schur functions is called the *Schur class*, indicated by $\mathcal{S}$.

The multipoint Schur algorithm goes as follows. Let (α_k), for $k \in \mathbb{N}$, be a *fixed* sequence of points in $\mathbb{D}$. We set $\alpha_0 = 0$ by convention. Define the elementary factor ζ_k by

$$\zeta_k(z) = \frac{z - \alpha_k}{1 - \bar{\alpha}_k z}, \tag{0.1}$$

and put for $f \in \mathcal{S}, k \geq 0$,

$$\begin{cases} f_0 = f, \\ \gamma_k = f_k(\alpha_{k+1}), \\ f_{k+1} = \dfrac{1}{\zeta_{k+1}} \dfrac{f_k - \gamma_k}{1 - \bar{\gamma}_k f_k}. \end{cases} \tag{0.2}$$

We call f_n the *Schur remainder* of f of order n. The *Schur convergent*, or *Schur approximant* to f of order n, is defined from (0.2) by formally computing f in terms of f_{n+1} and γ_k for $0 \leq k \leq n$, and then substituting $f_{n+1} = 0$ in the resulting expression.

It is a straightforward consequence of the maximum principle, that the algorithm stops at some finite n (*i.e.* that f_n is an unimodular constant) if and only if f is a Blaschke product of degree n, namely a rational function in $\mathcal{S}$ which is unimodular on $\mathbb{T}$:

$$B(z) = c \prod_{j=1}^{n} \frac{z - \beta_j}{1 - \bar{\beta}_j z},$$

where $\beta_j \in \mathbb{D}$, $|c| = 1$. *Throughout the paper, we assume that this is not the case*, so that the Schur algorithm, when applied to f with some sequence (α_k), produces an infinite sequence (f_k). By the maximum principle, it is easily seen that f_k is in turn Schur. The complex numbers γ_k appearing in the algorithm are called the *Schur (*or *Verblunsky) parameters* of f, and our assumption that f is not a finite Blaschke product is equivalent to the fact that $\gamma_k \in \mathbb{D}$ for all k.

The case where $\alpha_k \equiv 0$, originally considered by Schur [42] and subsequently studied by many authors, will be referred to as the *classical* Schur algorithm. *Thus, in the classical case,* $\alpha_k = 0$ *and* $\zeta_k(z) = z$ *for all* k*, as opposed to the multipoint version above where* (α_k) *may distribute arbitrarily in* $\mathbb{D}$.

It is clear from (0.2) (this is formalized in Proposition 1.2) that γ_k is completely determined by the interpolation values $f^{(j)}(\alpha_l)$ with $0 \leq j \leq n_l - 1$, where n_l is the multiplicity of α_l in the sequence $(\alpha_\ell)_{1\leq\ell\leq k+1}$ and the superscript (j) indicates the j-th derivative. In order for the Schur approximants to actually converge to f, it is thus necessary that the sequence

(α_k) be a uniqueness set in $H^\infty(\mathbb{D})$. This is equivalent to the negation of Blaschke condition:

$$\sum_k (1-|\alpha_k|) = +\infty. \tag{0.3}$$

Of importance to us will be the equivalence of (0.3) with the density of rational functions having poles at the points $(1/\overline{\alpha}_k)$ in every Hardy space $H^p(\mathbb{D})$, $1 \le p < \infty$, as well as in the disk algebra $A(\mathbb{D})$ [2, App. A].

Next, we recall a basic construction relating the classical Schur algorithm to orthogonal polynomials on $\mathbb{T}$, see *e.g.* [22, 43]. For μ a Borel probability measure on $\mathbb{T}$, we let μ_{ac} and μ_s respectively be its absolutely continuous and singular components with respect to m, the Lebesgue measure given by $dm(t) = dt/(2\pi i t) = \frac{1}{2\pi} d\theta$ where $t = e^{i\theta} \in \mathbb{T}$. We further put $\mu' = d\mu_{ac}/dm$ so that $d\mu = \mu' dm + d\mu_s$.

To $f \in \mathcal{S}$, we associate two probability measures $\mu, \tilde{\mu}$ on $\mathbb{T}$ by the relations

$$F_\mu(z) = \frac{1+zf}{1-zf} = \int_{\mathbb{T}} \frac{t+z}{t-z} d\mu(t), \quad F_{\tilde{\mu}}(z) = \frac{1-zf}{1+zf} = \int_{\mathbb{T}} \frac{t+z}{t-z} d\tilde{\mu}(t). \tag{0.4}$$

Clearly F_μ is a Carathéodory function, *i.e.* $\operatorname{Re} F(z) > 0,\ z \in \mathbb{D}$; moreover $F(0) = 1$. We call F_μ the *Herglotz transform* of μ, and the representation (0.4) is possible because every Carathéodory function is uniquely the Herglotz transform of a finite positive measure. From the Fatou theorems [25, Ch. I, Sect. D], we note that

$$\mu' = \operatorname{Re} F_\mu = \frac{1-|f|^2}{|1-zf|^2}, \qquad \lim_{r\to 1} \operatorname{Re} F_\mu(re^{i\theta}) = +\infty, \tag{0.5}$$

m-a.e. and μ_s-a.e., respectively. Similar considerations hold for $\tilde{\mu}$.

Let (ϕ_n) and (ψ_n) be the orthonormal polynomials with respect to μ and $\tilde{\mu}$:

$$\int_{\mathbb{T}} \phi_n \overline{\phi}_m d\mu = \delta_{nm}, \qquad \int_{\mathbb{T}} \psi_n \overline{\psi}_m d\tilde{\mu} = \delta_{nm}, \tag{0.6}$$

here δ_{nm} is the Kronecker symbol. Our assumption that f is not a finite Blaschke product means that μ and $\tilde{\mu}$ have infinite support, therefore ϕ_n, ψ_n have exact degree n. The sequences (ϕ_n) and (ψ_n) are called respectively the orthonormal polynomials of first and second kind associated with μ. Clearly ϕ_n and ψ_n are unique up to a multiplicative unimodular constant. We normalize them so that their respective leading coefficients k_n and k'_n are positive.

For a polynomial ϕ of degree n, put $\phi^*(z) = z^n \overline{\phi(1/\bar{z})}$. This is again a polynomial of degree n. Note that $k_n = \overline{\phi^*_n(0)}$. The coefficients

$$\tilde{\gamma}_n = \tilde{\gamma}_n(\mu) = -\frac{\overline{\phi_{n+1}(0)}}{k_{n+1}} \tag{0.7}$$

are called *the Geronimus parameters* associated with (ϕ_n) (or with μ).

The following remarkable theorem, named after Geronimus, was proven almost simultaneously by Geronimus [19] and Wall [45].

Theorem. *Let $f \in \mathcal{S}$. If $\alpha_k \equiv 0$, the Schur parameters and the Geronimus parameters coincide, i.e. $\gamma_n = \tilde{\gamma}_n,\ n \ge 0$.*

Since trading μ for $\tilde{\mu}$ is tantamount to change f into $-f$, a corollary is that $\tilde{\gamma}(\mu) = -\tilde{\gamma}(\tilde{\mu})$.

We turn to the multipoint version of Geronimus' theorem, which is due essentially to Bultheel et al. [10] although the first explicit statement is apparently in Langer-Lasarow [28]. For this, orthogonal polynomials need to be generalized into orthogonal rational functions whose construction we now explain. Define the "partial" Blaschke products $\mathcal{B}_k$ by

$$\mathcal{B}_0(z) = 1, \quad \mathcal{B}_k(z) = \mathcal{B}_{k-1}(z)\zeta_k(z), \tag{0.8}$$

where ζ_k is given by (0.1) and $k \geq 1$. The functions $\{\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_n\}$ span the space

$$\mathcal{L}_n = \left\{ \frac{p_n}{\pi_n} \ : \ \pi_n(z) = \prod_{k=1}^{n}(1 - \bar{\alpha}_k z), \ p_n \in \mathcal{P}_n \right\}, \tag{0.9}$$

where $\mathcal{P}_n$ stands for the space of algebraic polynomials of degree at most n. In the classical case, that is when $\alpha_k = 0$ for all k, $\mathcal{L}_n$ coincides with $\mathcal{P}_n$.

Given a function g, we introduce the parahermitian conjugate g_* defined by $g_*(z) = \overline{g(1/\bar{z})}$. Observe that $|g_*| = |g|$ on $\mathbb{T}$ and that $\zeta_{n*} = {\zeta_n}^{-1}$, $\mathcal{B}_{k*} = {\mathcal{B}_k}^{-1}$. For $g \in \mathcal{L}_n$, we set $g^* = \mathcal{B}_n f_*$; clearly, $g^* \in \mathcal{L}_n$. There is no notational discrepancy since in the classical case the star operation agrees with the definition we gave before. Put $\mathcal{B}_{n,i} = \prod_{k=i}^{n} \zeta_k$. Each $g \in \mathcal{L}_n$ can be uniquely decomposed in the form

$$g = a_n\mathcal{B}_n + a_{n-1}\mathcal{B}_{n-1} + \cdots + a_1\mathcal{B}_1 + a_0,$$

and then

$$g^* = \bar{a}_0\mathcal{B}_{n,1} + \bar{a}_1\mathcal{B}_{n,2} + \cdots + \bar{a}_{n-2}\mathcal{B}_{n,n-1} + \bar{a}_{n-1}\mathcal{B}_{n,n} + \bar{a}_n.$$

It is plain that $a_n = \overline{g^*(\alpha_n)}$ and $a_0 = g(\alpha_1)$.

Now, pick a Schur function f which is not a Blaschke product, denote its Herglotz measure by μ (0.4), and consider $\mathcal{L}_n$ as a subspace of $L^2(\mu)$. This is possible since μ has infinite support. Let $(\phi_k)_{0 \leq k \leq n}$ be an orthonormal basis for $\mathcal{L}_n$ such that $\phi_0 = 1$ and $\phi_k \in \mathcal{L}_k \setminus \mathcal{L}_{k-1}$. Such a basis is easily obtained on applying the Gram-Schmidt orthonormalization process to $\mathcal{B}_0, \mathcal{B}_1, \ldots, \mathcal{B}_n$. We customary write

$$\phi_n = \kappa_n\mathcal{B}_n + a_{n,n-1}\mathcal{B}_{n-1} + \ldots + a_{n,1}\mathcal{B}_1 + a_{n,0}\mathcal{B}_0, \tag{0.10}$$

where $\kappa_n = \overline{\phi_n^*(\alpha_n)}$.

Definition 0.1. *The functions (ϕ_k) are called the orthogonal rational functions of the first kind associated to (α_k) and μ.*

The (ψ_n) arising from embedding $\mathcal{L}_n$ to $L^2(\tilde{\mu})$ are called the orthogonal rational functions of the second kind. Clearly, the orthogonal rational functions $(\phi_n), (\psi_n)$ defined in (0.10) reduce to the orthonormal polynomials from (0.6) in the classical case.

Generically, the dependence on the nodes (α_k) and the measure μ will be omitted. The words "orthogonal rational function" will be abbreviated as ORF or OR-function.

The definition of the *Geronimus parameters* $(\tilde{\gamma}_k)$ for OR-functions is

$$\tilde{\gamma}_n = -\frac{\overline{\phi_n(\alpha_{n-1})}}{\overline{\phi_n^*(\alpha_{n-1})}}, \quad n \geq 1.$$

Note we do not define $\tilde{\gamma}_0$ and there is a shift of index as compared to (0.7).

It is quite nontrivial that one can relate the Schur algorithm (0.2) and the ORFs (0.10) in the multipoint case as well:

Theorem ([10, 28])**.** *Let (α_k), $f \in \mathcal{S}$, and the ORFs (ϕ_n) be as above. Then the multipoint Schur and Geronimus parameters coincide, i.e. $\gamma_k = \tilde{\gamma}_{k+1}$.*

We prove this fundamental result in Section 2 for the sake of completeness.

0.2. **Discussion of the main results.** The convergence properties of the Schur approximants and of the ORFs (ϕ_n) are the main address of the present work, which is in part inspired by the results obtained by Khrushchev [22]. To better see the parallel between the classical and the multipoint case, we give below a sample of results from [22] in the classical situation, and have them followed by their multipoint counterparts, numbered with a prime superscript; we connect these counterparts to the forthcoming results in between parentheses.

We say that a measure μ is Erdős, iff $\mu' > 0$ a.e. on $\mathbb{T}$. This is equivalent to say that $|f| < 1$ a.e. on $\mathbb{T}$.

Theorem 1 ([22], Theorem 1)**.** *Let $f \in \mathcal{S}$ and μ be its Herglotz measure. If $\alpha_k \equiv 0$, then μ is Erdős if and only if the Schur remainders f_n satisfy*

$$\lim_n \int_{\mathbb{T}} |f_n|^2 dm = 0.$$

The next result is stated in terms of the classical Wall polynomials A_n, B_n of f [22, Sect. 4], obtained from Definition 1.5 below by setting $\alpha_k \equiv 0$. By definition of the Wall polynomials, the ratio A_n/B_n is the *Schur approximant* to f of degree n. Recall that the pseudohyperbolic distance on $\mathbb{D}$ is defined as $\rho(z, w) = |z - w|/|1 - \bar{w}z|$, $z, w \in \mathbb{D}$.

Theorem 2 ([22], Corollary 2.4)**.** *A measure μ is Erdős if and only if*

$$\lim_n \int_{\mathbb{T}} \rho\left(f, \frac{A_n}{B_n}\right)^2 dm = 0.$$

We shall see that, in the multipoint situation when the sequence (α_k) accumulates on the unit circle, the conclusions of Theorems 1 and 2 get localized around the accumulation points of (α_k) on $\mathbb{T}$ so that L^2-norms get weighted by the Poisson kernel at α_{n+1}. This is why, somewhat reminiscently of the Fatou theorem, we put extra-conditions on μ, locally around such points, to derive convergence properties. Namely, let $Acc(\alpha_k) = \overline{(\alpha_k)} \backslash (\alpha_k)$ be the set of accumulation points of (α_k); the bar (or *clos* (.)) stands for the closure of a set. The following assumptions play an important role in our proofs

$$\mu' \in \mathcal{C}(\mathcal{O}(Acc(\alpha_k) \cap \mathbb{T})), \tag{0.11}$$

$$\mu' > 0 \ \text{ on } \ \mathcal{O}(Acc(\alpha_k) \cap \mathbb{T}), \tag{0.12}$$

$$\{Acc(\alpha_k) \cap \mathbb{T}\} \subset \mathbb{T} \backslash \operatorname{supp} \mu_s, \tag{0.13}$$

where, for $A \subset \mathbb{T}$, $\mathcal{O}(A)$ designates an open neighborhood of A in $\mathbb{T}$ and $\mathcal{C}(A)$ is the space of continuous functions on A. The *closed* support of μ_s is denoted by $\operatorname{supp}\mu_s$. When the sequence (α_k) accumulates nontangentially on $Acc(\alpha_k)\cap\mathbb{T}$, meaning that every convergent subsequence to $\xi\in\mathbb{T}$ tends to the latter nontangentially, two weaker substitutes for (0.11), (0.12) are also of interest:

(0.14) each $\xi \in Acc(\alpha_k)\cap\mathbb{T}$ is a Lebesgue point of μ', $\sqrt{\mu'}$, and $\mu'(\xi) > 0$;

(0.15) μ' is upper semicontinuous, $0<\delta<\mu'<M<\infty$ on $\mathcal{O}(Acc(\alpha_k)\cap\mathbb{T})$, and each $\xi\in Acc(\alpha_k)\cap\mathbb{T}$ is a Lebesgue point of $\log\mu'$.

From (0.5), we see that (0.11) and (0.12) may be ascertained in terms of f, namely $f\in\mathcal{C}(\mathcal{O}(Acc\,(\alpha_k)\cap\mathbb{T}))$ and $|f|<1$ there, while $(Acc(\alpha_k)\cap\mathbb{T})\subset\mathbb{T}\backslash clos\,\{z:\,zf(z)=1\}$.

The multipoint analogues of the previous theorems go as follows.

Theorem 1' (Corollary 3.5 and Theorem 4.6)**.** *Let* (0.3)*,* (0.11)-(0.13) *hold, and* $|f|<1$ *a.e. on* $\mathbb{T}$*. Then*

$$\lim_k \int |f_k|^2 P(.,\alpha_k)dm = 0.$$

If (α_k) *accumulates nontangentially on* $Acc(\alpha_k)\cap\mathbb{T}$*, then one can replace hypotheses* (0.11) *and* (0.12) *with* (0.14)*.*

Above, $P(.,\alpha_k)$ is the Poisson kernel at α_k on $\mathbb{D}$ (0.18). Denote by $W^{q,p}(\mathbb{T})$ the Sobolev spaces on $\mathbb{T}$ (see Section 0.3 for more details). Recall that $(A_n),(B_n)$ are the Wall rational functions from Definition 1.5 corresponding to $f\in\mathcal{S}$.

Theorem 1". *Let* μ *be absolutely continuous with* $\mu'\in W^{1-1/p,p}(\mathbb{T})$ *for some* $p>4$*, and* $\mu'>0$ *on some neighborhood* $\mathcal{O}(Acc(\alpha_k)\cap\mathbb{T})$*. Then*

$$\lim_n \left\| \left(f-\frac{A_n}{B_n}\right)\sqrt{P(.,\alpha_{n+1})}\right\|_\infty = 0.$$

Remarks on the converse to Theorem 1' follow Theorems 3.2, 3.4.

Theorem 2' (Theorem 4.2)**.** *Assumptions being as in Theorem 1', we have*

$$\lim_k \int_{\mathbb{T}} \rho\left(f,\frac{A_k}{B_k}\right)^2 P(.,\alpha_{k+1})dm = 0.$$

The final point of the paper is to carry over the Szegő theory to the multipoint setting. Recall that a measure μ is called *Szegő* (notation: $\mu\in$ (S)) iff $\log\mu'\in L^1(\mathbb{T})$. For $\mu\in$ (S), the associated Szegő function S is

$$S(z)=S[\mu](z):=\exp\left(\frac{1}{2}\int_{\mathbb{T}}\frac{t+z}{t-z}\log\mu' dm(t)\right). \tag{0.16}$$

The function S given by (0.16) is the so-called *outer* function in $H^2(\mathbb{D})$ such that $|S|^2=\mu'$ a.e. on $\mathbb{T}$, normalized so that $S(0)>0$.

The first version of the next theorem, which addresses the classical case, was proven by Szegő [44]. Subsequent improvement were obtained by Geronimus [20], Krein [26], and others; see Simon [43] for the discussion and a full list of references. Some of the latest improvements are due to Nikishin-Sorokin [36], Peherstorfer-Yuditskii [39]. A generalized version of Szegő condition is treated in Denisov-Kupin [14, 15].

Theorem 3. *Let* $\mu \in$ (S) *and* (ϕ_n) *be the corresponding orthonormal polynomials* (0.6)*. Then*

- $\lim_n (S\phi_n^*)(0) = 1$*; more generally,* $\lim_n (S\phi_n^*)(z) = 1$ *for* $z \in \mathbb{D}$*.*
- $\lim_n \int_{\mathbb{T}} |S\phi_n^* - 1|^2 dm = 0.$

Moreover, $\mu \in$ (S) *if and only if*

$$\lim_n \int_{\mathbb{T}} \mathfrak{P}\left(f, \frac{A_n}{B_n}\right)^2 dm = 0,$$

where $\mathfrak{P}(.,.)$ *is the hyperbolic distance on* $\mathbb{D}$ (4.1)*. Equivalently,*

$$\lim_n \int_{\mathbb{T}} \log(1 - |f_n|^2)\, dm = 0.$$

The last assertion of the theorem concerning the hyperbolic distance is from Khrushchev [22], Theorem 2.6.

A multipoint analogue to the previous theorem when (α_n) is compactly supported in $\mathbb{D}$ is Theorem 9.6.9 from Bultheel et al. [10]; its generalization to sequences (α_k) meeting (0.3) is given below. It is more difficult and requires some preparation. It relies on *a priori* pointwise estimates of (ϕ_n) (see Proposition 5.6), that play here the role of classical bounds by Szegő and Geronimus [44, Ch. 12], [20, Ch. 4]. Such estimates are new even in the polynomial case, as they handle some situations where μ' may vanish. Their proof in turn depends on $\overline{\partial}$-estimates and Sobolev embeddings. As compared to the case where (α_n) is compactly supported in $\mathbb{D}$, the result below is of new type in that $S\phi_n^*$ is asymptotic to a normalized Cauchy kernel at the last interpolation point, which is unbounded when (α_n) approaches $\mathbb{T}$. Here is a combination of Theorem 4.3, Theorem 5.8 and Corollary 5.13 to come:

Theorem 3'. *Let* (0.3)*,* (0.11)*-*(0.13) *be in force, with* $\mu \in$ (S)*. Then*

- $\lim_n |\phi_n^*(\alpha_n)|^2 |S(\alpha_n)|^2 (1 - |\alpha_n|^2) = 1$*; more generally, for any sequence* $(z_n) \subset \mathbb{D}$*, it holds*

$$\lim_n \left\{ \phi_n^*(z_n) S(z_n) \sqrt{1 - |z_n|^2} - \beta_n \frac{\sqrt{1 - |\alpha_n|^2}\sqrt{1 - |z_n|^2}}{1 - \overline{\alpha}_n z_n} \right\} = 0,$$

  where $\beta_n = (S\phi_n^*)(\alpha_n)/|(S\phi_n^*)(\alpha_n)|$*. In particular, for a fixed* $z \in \mathbb{D}$*,*

$$\lim_n \left\{ S\phi_n^*(z) - \beta_n \frac{\sqrt{1 - |\alpha_n|^2}}{1 - \overline{\alpha}_n z} \right\} = 0.$$

- *We also have*

$$\lim_n \left\| S\phi_n^*(z) - \beta_n \frac{\sqrt{1 - |\alpha_n|^2}}{1 - \overline{\alpha}_n z} \right\|_2 = 0.$$

- *Moreover*

$$\lim_n \int_{\mathbb{T}} \mathfrak{P}\left(f, \frac{A_n}{B_n}\right)^2 P(., \alpha_{n+1}) dm = 0,$$

in particular

$$\lim_n \int_{\mathbb{T}} \log(1 - |f_n|^2) P(., \alpha_n) dm = 0.$$

If (α_k) *accumulates nontangentially on* $Acc(\alpha_k) \cap \mathbb{T}$*, then one can replace* (0.11) *and* (0.12) *by* (0.15).

It would be interesting to know how much these assumptions can be relaxed. In particular, we shall give an example where the conclusion of Theorem 3' holds although (0.12) fails.

The paper is organized as follows. The multipoint Schur algorithm, its connections to continued fractions, the Wall rational functions and the Schur parameters are discussed in Section 1. Section 2 introduces the ORFs $(\phi_n), (\psi_n)$, and expresses them through Geronimus parameters and transfer matrices. The construction is then used to prove Geronimus' theorem and its corollaries. Although we use a different normalization, by and large, the content of Section 2 is borrowed from Bultheel et al. [10]. The convergence of Schur remainders and Wall RFs is studied in Sections 3 and 4. Section 5 is devoted to the discussion of the Szegő-type theorem and its corollaries.

0.3. **Some notation.** As already mentioned, the closure of $A \subset \mathbb{C}$ is indicated by $clos\, A$ or $\overline{A}$, while $\mathcal{O}(A)$ designates an open neighborhood of A in $\mathbb{T}$. The normalized Lebesgue measure on $\mathbb{T}$ is denoted by m, and the measure of $A \subset \mathbb{T}$ is denoted by $|A|$. We put $\mathcal{C}(A)$ for the space of continuous functions on A. The symbol $||.||_p$ stands for the usual norm on the Lebesgue space $L^p(\mathbb{T})$, $1 \leq p \leq \infty$; *when* $p = 2$, *the subindex is usually dropped.* The classical analytic Hardy spaces of the disk are denoted by $H^p(\mathbb{D}), 1 \leq p \leq \infty$, and $A(\mathbb{D})$ is the disk algebra, comprised of analytic functions in $\mathbb{D}$ that extend continuously to $\overline{\mathbb{D}}$, endowed with the *sup* norm. Standard references on the subject are the books by Duren [16], Garnett [18], Koosis [25], from which we often quote basic facts without further citation. In particular, H^p-functions have well-defined nontangential limits in $L^p(\mathbb{T})$, and we use the same notation for the function in $\mathbb{D}$ and its trace on $\mathbb{T}$.

Every real-valued $\varphi \in L^1(\mathbb{T})$ is m-a.e. the real part of the nontangential limit of the complex analytic function

$$F_\varphi(z) = \int_{\mathbb{T}} \frac{t+z}{t-z} \varphi(t)\, dt, \qquad z \in \mathbb{D}, \tag{0.17}$$

which is called the *Herglotz transform* of φ. The map sending φ to the imaginary part of F_φ is the *conjugation operator*, denoted with the superscript "$\check{}$", *i.e.* $F_\varphi = \varphi + i\check{\varphi}$; it extends linearly to complex-valued functions. By a theorem of M. Riesz, the conjugation operator acts on $L^p(\mathbb{T}), 1 < p < \infty$. Moreover, since it is of convolution type, it commutes with $d/|dt|$ and integrating by parts one sees that it also acts on $W^{1,p}(\mathbb{T})$, the space of absolutely continuous functions with L^p derivative on $\mathbb{T}$.

The Poisson kernel on $\mathbb{D}$ is

$$P(z,w) = P_w(z) = (1-|w|^2)/|z-w|^2, \tag{0.18}$$

where $z \in \mathbb{T}, w \in \mathbb{D}$.

We shall need some basic facts from Sobolev space theory for which we refer the reader to Adams-Fournier [1]. In particular, for $I \subset \mathbb{T}$ an open arc, those φ for which

$$\int_{t,t' \in I} \left| \frac{\varphi(t) - \varphi(t')}{t-t'} \right|^p dm(t)dm(t') < \infty, \tag{0.19}$$

with $1 < p < \infty$, form the *fractional Sobolev space* $W^{1-1/p,p}(I)$ (that coincides with the Besov space $B_p^{1-1/p,p}(I)$). It is a real interpolation space between $W^{1,p}(I)$ and $L^p(I)$: $W^{1-1/p,p}(I) = [W^{1,p}(I), L^p(I)]_{1/p}$, that embeds compactly into $L^p(I)$. By interpolation [1, Sect. 7.3.2, Theorem 7.3.2], the conjugation operator also acts on $W^{1-1/p,p}(\mathbb{T})$.

When h is defined on $E \subset \mathbb{C}$ and $0 < \alpha < 1$, we say that h is Hölder continuous of exponent α if there is a constant $C > 0$ such that $|h(t) - h(t')| \leq C|t-t'|^\alpha$ for all $t, t' \in E$. We then write $h \in H_\alpha(E)$. Note that $H_\alpha(I) \subset W^{1-1/p,p}(I)$ if $1 < p < 1/(1-\alpha)$. By the Sobolev embedding theorem [1, Theorem 4.12], it holds conversely that $W^{1-1/p,p}(I) \subset H_{1-2/p}(I)$ for $p > 2$.

1. Wall rational functions

In this section we rewrite Section 4 from [22] for the multipoint case. The presentation is very close to the original, and only some technical details are different. Roughly speaking, one mainly has to replace z^n with $\mathcal{B}_n$; that is why we generically give results accompanied by precise references to [22] and omit the proofs. We start recalling basic definitions on continued fractions [46].

A continued fraction is an infinite expression of the form

$$b_0 + \cfrac{a_1}{b_1 + \cfrac{a_2}{b_2 + \cfrac{a_3}{b_3 + \frac{a_4}{\cdots}}}}.$$

We conform the more economic notation

$$b_0 + \frac{a_1}{b_1} + \frac{a_2}{b_2} + \frac{a_3}{b_3} + \cdots.$$

For any complex-valued ω, we let $t_0(\omega) = b_0 + \omega$ and, for $k \geq 1$,

$$t_k(\omega) = \frac{a_k}{b_k + \omega}.$$

By definition, the n-th convergent P_n/Q_n of the continued fraction is

$$\frac{P_n}{Q_n} = t_0 \circ t_1 \circ \cdots \circ t_n(0) = b_0 + \frac{a_1}{b_1} + \frac{a_2}{b_2} + \cdots + \frac{a_n}{b_n}.$$

Proposition 1.1 ([22], relations (3.2)-(3.4))**.** *The quantities P_n and Q_n can be computed according to the recurrence relations:*

$$\begin{cases} P_{-1} = 1, Q_{-1} = 0, \\ P_0 = b_0, Q_0 = 1, \\ P_{k+1} = b_{k+1}P_k + a_{k+1}P_{k-1} \\ Q_{k+1} = b_{k+1}Q_k + a_{k+1}Q_{k-1} \end{cases}$$

for $k \geq 0$. More generally,

$$\frac{P_{n-1}\omega + P_n}{Q_{n-1}\omega + Q_n} = t_0 \circ t_1 \circ \cdots \circ t_n(\omega).$$

First we record the following fact.

Proposition 1.2. *For $k \geq 1$, γ_k depends only on $f^{(i)}(\alpha_j)$, $1 \leq j \leq k+1$, $0 \leq i < m_j$, where m_j is the multiplicity of α_j at the k-th step, i.e. m_j is the number of times the value of α_j enters $(\alpha_l)_{1 \leq l \leq k+1}$.*

Proof. Noticing in case of repetitions that $f_j(\alpha_j) = f'_{j-1}(\alpha_j)\frac{1-|\alpha_j|^2}{1-|f_{j-1}(\alpha_j)|^2}$, the proof is immediate by induction on (0.2). □

We now rewrite the recursive step of (0.2) as

$$f_{k-1} = \gamma_{k-1} + \frac{(1-|\gamma_{k-1}|^2)\zeta_k}{\bar{\gamma}_{k-1}\zeta_k + \frac{1}{f_k}}. \tag{1.1}$$

For $\omega \in \mathbb{D} \setminus \{0\}$, set

$$\tau_k(\omega) = \tau_k(\omega, z) := \gamma_k + \frac{(1-|\gamma_k|^2)\zeta_{k+1}}{\bar{\gamma}_k\zeta_{k+1} + \frac{1}{\omega}}, \tag{1.2}$$

and put $\tau_k(0) = \gamma_k$. Hence, $f_k = \tau_k(f_{k+1})$ and

$$f = \tau_0 \circ \tau_1 \circ \cdots \circ \tau_n(f_{n+1}). \tag{1.3}$$

In a way reminiscent of how we defined P_n/Q_n, we obtain the Schur convergent R_n of degree n upon replacing f_{n+1} by 0 in (1.3), that is,

$$R_n = \tau_0 \circ \tau_1 \circ \cdots \circ \tau_{n-1} \circ \tau_n(0) = \tau_0 \circ \tau_1 \circ \cdots \circ \tau_{n-1}(\gamma_n). \tag{1.4}$$

Proposition 1.3. *The rational function R_n interpolates f at $(\alpha_k)_{1 \leq k \leq n+1}$, counting multiplicities, and their first $n+1$ Schur parameters coincide.*

Proof. Note that $\tau_k(\omega, \alpha_{k+1}) = \gamma_k$ is independent of ω. Thus, for $0 \leq k \leq n$,

$$\begin{aligned} f(\alpha_{k+1}) &= \tau_0 \circ \cdots \circ \tau_k(\tau_{k+1} \circ \cdots \circ \tau_n(f_{n+1}), \alpha_{k+1}) \\ &= \tau_0 \circ \cdots \circ \tau_k(\tau_{k+1} \circ \cdots \circ \tau_n(0), \alpha_{k+1}) \\ &= R_n(\alpha_{k+1}). \end{aligned}$$

Consequently, R_n interpolates f at the point α_{k+1}.

The remaining part of the claim is proven by induction. The base of induction being obvious, suppose that the k first Schur parameters of f and R_n coincide. Then, denoting $R_n^{[1]}, \ldots R_n^{[n]}$ the Schur remainders of R_n, we see that $R_n^{[k]} = \tau_{k-1}^{-1} \circ \cdots \circ \tau_0^{-1}(R_n)$, and

$$R_n^{[k]}(\alpha_{k+1}) = \tau_{k-1}^{-1} \circ \cdots \circ \tau_0^{-1}(R_n, \alpha_{k+1})$$

$$
\begin{aligned}
&= \tau_{k-1}^{-1} \circ \cdots \circ \tau_0^{-1} \circ \tau_0 \circ \tau_1 \circ \cdots \circ \tau_{n-1}(\gamma_n, \alpha_{k+1}) \\
&= \tau_k(\tau_{k+1} \circ \cdots \circ \tau_{n-1}(\gamma_n), \alpha_{k+1}) = \gamma_{k+1}.
\end{aligned}
$$

Therefore, the $k+1$-th Schur parameter of R_n and f coincide. $\square$

The Schur algorithm can be readily connected to the continued fractions. Indeed, let P_n/Q_n be the sequence of convergents associated to

$$
\gamma_0 + \frac{(1-|\gamma_0|^2)\zeta_1}{\bar\gamma_0\zeta_1} \ + \ \frac{1}{\gamma_1} \ + \ \frac{(1-|\gamma_1|^2)\zeta_2}{\bar\gamma_1\zeta_2} \ + \ldots. \tag{1.5}
$$

Then, the functions R_n are none but the P_{2n}/Q_{2n}.

For $n \geq 1$, we have by Proposition 1.1

$$
\begin{aligned}
P_{2n} &= \gamma_n P_{2n-1} + P_{2n-2} \\
Q_{2n} &= \gamma_n Q_{2n-1} + Q_{2n-2} \\
P_{2n-1} &= \bar\gamma_{n-1}\zeta_n P_{2n-2} + (1-|\gamma_{n-1}|^2)\zeta_n P_{2n-3} \\
Q_{2n-1} &= \bar\gamma_{n-1}\zeta_n Q_{2n-2} + (1-|\gamma_{n-1}|^2)\zeta_n Q_{2n-3}
\end{aligned} \tag{1.6}
$$

with

$$
P_{-1} = 1, \quad P_0 = \gamma_0, \quad Q_{-1} = 0, \quad Q_0 = 1.
$$

For P_{2n} and Q_{2n}, we easily prove the next lemma.

Lemma 1.4 ([22], Lemma 4.1)**.** *For $n \geq 0$, we have $P_{2n+1}, Q_{2n+1} \in \mathcal{L}_{n+1}$, $P_{2n}, Q_{2n} \in \mathcal{L}_n$ and*

$$
P_{2n+1} = \zeta_{n+1} Q_{2n}^*, \quad Q_{2n+1} = \zeta_{n+1} P_{2n}^* \ .
$$

Mimicking [22], formulas (4.5), (4.12), we get

$$
\begin{bmatrix} Q_{2n}^* & P_{2n}^* \\ P_{2n} & Q_{2n} \end{bmatrix} = \left(\prod_{k=n}^{1} \begin{bmatrix} 1 & \bar\gamma_k \\ \gamma_k & 1 \end{bmatrix} \begin{bmatrix} \zeta_k & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 & \bar\gamma_0 \\ \gamma_0 & 1 \end{bmatrix}, \tag{1.7}
$$

with $n \geq 1$. Let us set $A_n = P_{2n}$, $B_n = Q_{2n}$ and choose the representative $R_n = A_n/B_n$ for R_n.

Definition 1.5. *A_n and B_n are called the n-th* W*all rational functions associated to the Schur function f and the sequence (α_k).*

In the new notation, the previous relation reads as

Proposition 1.6 ([22], relation (4.12))**.** *We have*

$$
\begin{bmatrix} B_n^* & A_n^* \\ A_n & B_n \end{bmatrix} = \left(\prod_{k=n}^{1} \begin{bmatrix} 1 & \bar\gamma_k \\ \gamma_k & 1 \end{bmatrix} \begin{bmatrix} \zeta_k & 0 \\ 0 & 1 \end{bmatrix} \right) \begin{bmatrix} 1 & \bar\gamma_0 \\ \gamma_0 & 1 \end{bmatrix}. \tag{1.8}
$$

The dependence of A_n, B_n on f and (α_k) will be usually dropped. For convenience, we abbreviate "Wall rational function" as WRF or Wall RF.

Corollary 1.7 ([22], relations (4.14), (4.15))**.** *The Wall RFs A_n, B_n have the following properties:*

1. $B_n(z)B_n^*(z) - A_n(z)A_n^*(z) = \mathcal{B}_n(z)\omega_n$,
2. $|B_n(\xi)|^2 - |A_n(\xi)|^2 = \omega_n$ *on* $\mathbb{T}$,
3. $f(\alpha_i) = A_n/B_n(\alpha_i) = B_n^*/A_n^*(\alpha_i)$, *for* $1 \leq i \leq n+1$,

where

$$\omega_n = \prod_{k=0}^{n} (1 - |\gamma_k|^2).$$

The proof is by taking the determinant in (1.8).

Proposition 1.8 ([22], Lemma 4.5). *For $n \geq 0$, we have*

(1) A_n, A_n^* *and* B_n *lie in* $\mathcal{L}_n$,
(2) B_n *does not vanish on* $\overline{\mathbb{D}}$,
(3) A_n/B_n *and* A_n^*/B_n *are Schur functions.*

These preparations bring us to the following important

Theorem 1.9 ([22], Theorem 4.6). *The Wall RFs A_n and B_n are connected to f and f_{n+1} by*

$$f(z) = \frac{A_n(z) + \zeta_{n+1}(z) B_n^*(z) f_{n+1}(z)}{B_n(z) + \zeta_{n+1}(z) A_n^*(z) f_{n+1}(z)}. \tag{1.9}$$

The theorem shows that, in Nevanlinna's parametrization of all Schur interpolants to f at $(\alpha_k)_{1 \leq k \leq n+1}$ [18, Ch. IV, Lemma 6.1], the value zero for the parameter yields $R_n = A_n/B_n$ while the value f_{n+1} yields f.

2. ORFs and Geronimus' theorem

The results of this section are borrowed from [10, 11]. We formulate the results and briefly discuss them for the completeness of presentation; the proofs are generically omitted.

2.1. **Orthogonal rational functions.** Let μ be a positive probability measure on $\mathbb{T}$ with infinite support. Obviously, $\mathcal{L}_n$ is a (closed) subspace of $L^2(\mu)$, and, following [10, Ch. 3], we regard it as a reproducing kernel Hilbert space. The reproducing kernels for $\mathcal{L}_n$ are easily seen to satisfy the so-called Christoffel-Darboux relations, which can be interpreted as recurrence relations for the ORFs (ϕ_n). Namely, we have

Theorem 2.1 ([10], Theorem 4.1.1). *For $n \geq 1$, it holds that*

$$\begin{bmatrix} \phi_n(z) \\ \phi_n^*(z) \end{bmatrix} = T_n(z) \begin{bmatrix} \phi_{n-1}(z) \\ \phi_{n-1}^*(z) \end{bmatrix},$$

where

$$\begin{aligned} T_n(z) \;&=\; \sqrt{\frac{1-|\alpha_n|^2}{1-|\alpha_{n-1}|^2}} \frac{1}{\sqrt{1-|\tilde{\gamma}_n|^2}} \frac{1-\bar{\alpha}_{n-1}z}{1-\bar{\alpha}_n z} \begin{bmatrix} 1 & -\overline{\tilde{\gamma}_n} \\ -\tilde{\gamma}_n & 1 \end{bmatrix} \\ &\times \begin{bmatrix} \lambda_n & 0 \\ 0 & \bar{\lambda}_n \end{bmatrix} \begin{bmatrix} \zeta_{n-1}(z) & 0 \\ 0 & 1 \end{bmatrix}, \end{aligned} \tag{2.1}$$

and

$$\tilde{\gamma}_n = -\frac{\overline{\phi_n(\alpha_{n-1})}}{\overline{\phi_n^*(\alpha_{n-1})}}, \qquad \eta_n = \frac{1-\alpha_n\bar{\alpha}_{n-1}}{1-\bar{\alpha}_n\alpha_{n-1}}, \tag{2.2}$$

$$\lambda_n = \frac{|1-\bar{\alpha}_n\alpha_{n-1}|}{1-\alpha_n\bar{\alpha}_{n-1}} \frac{\overline{\phi_n^*(\alpha_{n-1})}}{|\phi_n^*(\alpha_{n-1})|} \frac{\overline{\kappa_{n-1}}}{|\kappa_{n-1}|} \eta_n. \tag{2.3}$$

Definition 2.2. *We call $\tilde{\gamma}_n$, given by* (2.2)*, the n-th Geronimus parameter of the measure μ (with respect to the sequence (α_k)).*

Corollary 3.1.4 from [10] says that for $z \in \mathbb{D}$, $n \geq 1$

$$\phi_n^*(z) \neq 0, \qquad |\phi_n(z)/\phi_n^*(z)| < 1, \tag{2.4}$$

and, consequently, $\tilde{\gamma}_n$ is well-defined and that $|\tilde{\gamma}_n| < 1$.

We will normalize ϕ_n by setting $\lambda_n = 1$, see (2.3). *Thus from now on, ϕ_n is the orthogonal rational function of degree n satisfying:*

$$\lambda_n = \frac{1 - \alpha_n \bar{\alpha}_{n-1}}{|1 - \alpha_n \bar{\alpha}_{n-1}|} \frac{\overline{\phi_n^*(\alpha_{n-1})}}{|\phi_n^*(\alpha_{n-1})|} \frac{\overline{\kappa_{n-1}}}{|\kappa_{n-1}|} = 1. \tag{2.5}$$

This normalization is from Langer-Lasarow [28]. It differs from the one made in Bultheel et al. [10], that corresponds to $\kappa_n = \overline{\phi_n^*(\alpha_n)} > 0$. However, in the classical case, $\alpha_n \equiv 0$, it is easily checked by induction, that it also matches the normalization $k_n > 0$ made in (0.7).

Relations (2.4) mean that the roots of ϕ_n lie in $\overline{\mathbb{D}}$. Theorem 2.1 implies that the roots of the orthogonal rational functions ϕ_n are, in fact, in $\mathbb{D}$. Another useful fact is that the OR-functions $(\phi_k)_{0 \leq k \leq n}$, are orthonormal in $L^2\left(\frac{P(.,\alpha_n)}{|\phi_n|^2} dm\right)$, see [10, Theorem 6.1.9].

We already saw a definition of ORFs of the second kind (see the discussion following Definition 0.1). Presently, the OR-functions of the second kind will be introduced by an explicit formula:

$$\begin{cases} \psi_0 = 1, \\ \psi_n(z) = \displaystyle\int_{\mathbb{T}} \frac{t+z}{t-z} \left(\phi_n(t) - \phi_n(z)\right) d\mu(t). \end{cases} \tag{2.6}$$

Both definitions turn out to be equivalent (see Theorem 2.6 or [10, Theorem 6.2.5]), but the one above is better suited for computations. The next result wraps Lemmas 4.2.2 and 4.2.3 from [10], whose proof is a direct computation using the orthogonality of (ϕ_n).

Lemma 2.3. *Let $n \geq 1$ and the function g be so that $g_* \in \mathcal{L}_{n-1}$. Then*

$$\psi_n(z) g(z) = \int_{\mathbb{T}} \frac{t+z}{t-z} \left(\phi_n(t) g(t) - \phi_n(z) g(z)\right) d\mu(t).$$

Similarly, for h such that $h_* \in \zeta_n \mathcal{L}_{n-1}$, we have

$$-\psi_n^*(z) h(z) = \int_{\mathbb{T}} \frac{t+z}{t-z} \left(\phi_n^*(t) h(t) - \phi_n^*(z) h(z)\right) d\mu(t). \tag{2.7}$$

Recall the Herglotz transform F_μ of a measure μ defined in (0.4). Plugging $h = (\mathcal{B}_n)_*$ in (2.7) and using that ϕ_n is μ-orthogonal to constants, we obtain at once

Proposition 2.4 ([11], Theorem 3.4)**.** *Let ϕ_n be the ORF of the first kind and ψ_n be as in* (2.6)*. Then*

$$F_\mu(z) = \frac{\psi_n^*(z)}{\phi_n^*(z)} + \frac{z \mathcal{B}_n(z) u_n(z)}{\phi_n^*(z)}, \tag{2.8}$$

where u_n is a analytic function in $\mathbb{D}$ given by

$$u_n(z) = 2\int_{\mathbb{T}} (\phi_n)_*(t)\,\frac{d\mu(t)}{t-z}, \qquad z \in \mathbb{D}. \tag{2.9}$$

In particular, ψ_n^*/ϕ_n^* interpolates F_μ at 0 and at the α_k for $1 \le k \le n$.

The theorem to come is [10, Theorem 4.2.4], with a different normalization.

Theorem 2.5. *The ORFs (ϕ_n) and the (ψ_n) from (2.6) together satisfy:*
(2.10)

$$\begin{bmatrix} \phi_n & \psi_n \\ \phi_n^* & -\psi_n^* \end{bmatrix} = \frac{\sqrt{1-|\alpha_n|^2}}{1-\bar{\alpha}_n z}\frac{1}{\Pi_n}\left(\prod_{k=n}^{1}\begin{bmatrix} 1 & -\overline{\tilde{\gamma}_k} \\ -\tilde{\gamma}_k & 1 \end{bmatrix}\begin{bmatrix} \zeta_{k-1}(z) & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

where $\Pi_n = \prod_{k=n}^{k=1}\sqrt{1-|\tilde{\gamma}_k|^2}$. In particular, ψ_n is in $\mathcal{L}_n$.

By taking determinants in (2.10), we get for $z \in \mathbb{D}$

$$\phi_n(z)\psi_n^*(z) + \phi_n^*(z)\psi_n(z) = 2\frac{1-|\alpha_n|^2}{(1-\overline{\alpha_n}z)(z-\alpha_n)}z\mathcal{B}_n(z),$$

and, consequently, for $z \in \mathbb{T}$,

$$\phi_n(z)\psi_n^*(z) + \phi_n^*(z)\psi_n(z) = 2\mathcal{B}_n(z)P(z,\alpha_n). \tag{2.11}$$

2.2. **Geronimus theorem.** The Geronimus-type theorem below is central for the whole construction. It seems first stated in [28], but it is implicitly contained in [10, Sect. 6.4].

Theorem 2.6. *Let $f \in \mathcal{S}$ and μ be the measure associated to f by (0.4). Then, for $k \ge 0$,*

$$\tilde{\gamma}_{k+1} = \gamma_k,$$

where $(\tilde{\gamma}_k)$ are the Geronimus parameters defined in (2.2) and (γ_k) the Schur parameters defined in (0.2).

Thus the Geronimus parameters and the Schur parameters of of a measure μ coincide. It follows from (2.10) that (ψ_n) meets the same recurrence relations as (ϕ_n) only with Geronimus parameters $-\tilde{\gamma}_n$ rather than $\tilde{\gamma}_n$. Thus, we see that the definition of the ORFs of the second kind given in (2.6) coincides with the one made in the introduction.

Proof. The idea is to compare the recurrence formulas (1.8) and (2.10). We assume the sequence (α_k) is simple, *i.e.* $\alpha_k \ne \alpha_j$ for $k \ne j$. The proof in the general case follows by a limiting argument. By (2.10), we have

$$\begin{aligned}&\begin{bmatrix} \phi_{n+1}(z) & \psi_{n+1}(z) \\ \phi_{n+1}^*(z) & -\psi_{n+1}^*(z) \end{bmatrix} \\ &= \Delta_{n+1}\left(\prod_{k=n+1}^{k=1}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} 1 & \overline{\tilde{\gamma}_k} \\ \tilde{\gamma}_k & 1 \end{bmatrix}\begin{bmatrix} \zeta_{k-1}(z) & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\right)\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\end{aligned}$$

where

$$\Delta_{n+1} = \frac{\sqrt{1-|\alpha_{n+1}|^2}}{1-\bar{\alpha}_{n+1}z}\frac{1}{\prod_{k=1}^{n+1}\sqrt{1-|\tilde{\gamma}_k|^2}}.$$

Let now U_n/V_n be the n-th convergent of the Schur function with Schur parameters $\gamma_k := \tilde{\gamma}_{k+1},\ k \geq 0$. Proposition 1.6 provides us with the following expression for ϕ_n, ψ_n:

$$\begin{aligned}
&\begin{bmatrix} \phi_{n+1}(z) & \psi_{n+1}(z) \\ \phi^*_{n+1}(z) & -\psi^*_{n+1}(z) \end{bmatrix} \\
&= \Delta_{n+1} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_n^* & U_n^* \\ U_n & V_n \end{bmatrix} \begin{bmatrix} \zeta_0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \\
&= \Delta_{n+1} \begin{bmatrix} zV_n^* - U_n^* & zV_n^* + U_n^* \\ -zU_n + V_n & -zU_n - V_n \end{bmatrix}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\begin{bmatrix} \phi_{n+1}(z) & \psi_{n+1}(z) \\ \phi^*_{n+1}(z) & -\psi^*_{n+1}(z) \end{bmatrix} \\
&= \frac{\sqrt{1-|\alpha_{n+1}|^2}}{1-\bar{\alpha}_{n+1}z} \frac{1}{\prod_{k=1}^{n+1}\sqrt{1-|\tilde{\gamma}_k|^2}} \begin{bmatrix} zV_n^* - U_n^* & zV_n^* + U_n^* \\ -zU_n + V_n & -zU_n - V_n \end{bmatrix},
\end{aligned} \tag{2.12}$$

and

$$\frac{\psi^*_{n+1}}{\phi^*_{n+1}} = \frac{1 + z\frac{U_n}{V_n}}{1 - z\frac{U_n}{V_n}}. \tag{2.13}$$

Consequently,

$$\frac{U_n(z)}{V_n(z)} = \Omega_z\left(\frac{\psi^*_{n+1}(z)}{\phi^*_{n+1}(z)}\right),$$

where $\Omega_z(w) = (w-1)/(z(w+1))$. From Proposition 2.4, we get

$$F(\alpha_{j+1}) = \left(\frac{\psi^*_{n+1}}{\phi^*_{n+1}}\right)(\alpha_{j+1}).$$

Recalling that $f(z) = \Omega_z(F(z))$, it follows by Proposition 1.2 that the $n+1$ first Schur parameters of the function U_n/V_n and of the function f coincide. $\square$

The theorem shows that the functions U_n and V_n are equal to the WRFs A_n and B_n corresponding to f. In particular, (2.12) and (2.13) imply

$$\begin{aligned}
&\begin{bmatrix} \phi_{n+1}(z) & \psi_{n+1}(z) \\ \phi^*_{n+1}(z) & -\psi^*_{n+1}(z) \end{bmatrix} \\
&= \frac{\sqrt{1-|\alpha_{n+1}|^2}}{1-\bar{\alpha}_{n+1}z} \frac{1}{\prod_{k=1}^{n+1}\sqrt{1-|\tilde{\gamma}_k|^2}} \begin{bmatrix} zB_n^* - A_n^* & zB_n^* + A_n^* \\ -zA_n + B_n & -zA_n - B_n \end{bmatrix}
\end{aligned} \tag{2.14}$$

and

$$\frac{\psi^*_{n+1}}{\phi^*_{n+1}} = \frac{1 + z\frac{A_n}{B_n}}{1 - z\frac{A_n}{B_n}}. \tag{2.15}$$

2.3. **Consequences of Geronimus theorem.** Arguing as in [22, Corollary 5.2], we readily see that the Schur function A_n/B_n corresponds to the measure $\frac{P(.,\alpha_{n+1})}{|\phi_{n+1}|^2}dm$.

The next theorem provides one with a helpful relation between the density μ' of the absolutely continuous part of μ, the Schur remainders (f_n), and the ORFs (ϕ_n). It is a counterpart to Theorem 2 from [22], see also [37]. Since it is heavily used in the sequel, we give the proof.

Theorem 2.7. *Let (ϕ_n) and (f_n) be the ORFs and Schur remainders associated to μ and f, respectively. Then it holds a.e. on $\mathbb{T}$ that*

$$\mu' = \frac{1-|f_n|^2}{|1-\zeta_n\frac{\phi_n}{\phi_n^*}f_n|^2}\frac{P(.,\alpha_n)}{|\phi_n|^2}.$$

Proof. From Theorem 1.9, we have on $\mathbb{T}$

$$\begin{aligned} 1-|f|^2 &= 1-\left|\frac{A_n+\zeta_{n+1}B_n^*f_{n+1}}{B_n+\zeta_{n+1}A_n^*f_{n+1}}\right|^2 \\ &= \frac{|B_n+\zeta_{n+1}A_n^*f_{n+1}|^2-|A_n+\zeta_{n+1}B_n^*f_{n+1}|^2}{|B_n+\zeta_{n+1}A_n^*f_{n+1}|^2}. \end{aligned} \tag{2.16}$$

Notice that $A_n^*\overline{B_n}=\overline{A_n}B_n^*$ on $\mathbb{T}$, so that

$$\zeta_{n+1}A_n^*f_{n+1}\overline{B_n}+B_n\overline{\zeta_{n+1}A_n^*f_{n+1}}-\overline{A_n}\zeta_{n+1}B_n^*f_{n+1}-A_n\overline{\zeta_{n+1}B_n^*f_{n+1}}=0.$$

Therefore, on expanding (2.16) and recalling Corollary 1.7, we find that

$$1-|f|^2=\frac{(|B_n|^2-|A_n|^2)(1-|f_{n+1}|^2)}{|B_n+\zeta_{n+1}A_n^*f_{n+1}|^2}=\frac{\omega_n(1-|f_{n+1}|^2)}{|B_n+\zeta_{n+1}A_n^*f_{n+1}|^2},$$

where $\omega_n=\prod_{k=0}^n(1-|\gamma_k|^2)$.

Again by Theorem 1.9, we obtain

$$\begin{aligned} |1-zf|^2 &= \left|1-\frac{zA_n+\zeta_{n+1}zB_n^*f_{n+1}}{B_n+\zeta_{n+1}A_n^*f_{n+1}}\right|^2 \\ &= \left|\frac{B_n-zA_n+\zeta_{n+1}f_{n+1}(A_n^*-zB_n^*)}{B_n+\zeta_{n+1}A_n^*f_{n+1}}\right|^2. \end{aligned}$$

On the other hand, Theorem 2.6 and (2.14) show

$$\begin{cases} zB_n^*-A_n^* &= \frac{1-\bar\alpha_{n+1}z}{\sqrt{1-|\alpha_{n+1}|^2}}\sqrt{\omega_n}\phi_{n+1} \\ B_n-zA_n &= \frac{1-\bar\alpha_{n+1}z}{\sqrt{1-|\alpha_{n+1}|^2}}\sqrt{\omega_n}\phi_{n+1}^* \end{cases}$$

and therefore

$$|1-zf|^2=\omega_n\frac{|1-\bar\alpha_{n+1}z|^2}{1-|\alpha_{n+1}|^2}\left|\frac{\phi_{n+1}^*-\zeta_{n+1}f_{n+1}\phi_{n+1}}{B_n+\zeta_{n+1}A_n^*f_{n+1}}\right|^2.$$

Recall that $\mu'(\xi)=(1-|f(\xi)|^2)/(|1-\xi f(\xi)|^2)$ a.e. on $\mathbb{T}$. Combining all this, we obtain

$$\mu'=\frac{1-|f_{n+1}|^2}{|\phi_{n+1}|^2|1-\zeta_{n+1}\frac{\phi_{n+1}}{\phi_{n+1}^*}f_{n+1}|^2}\frac{1-|\alpha_{n+1}|^2}{|\xi-\alpha_{n+1}|^2}$$

which achieves the proof. □

3. Weighted L^2-convergence of Schur remainders

The material reviewed so far is known, and it is meant as a preparation for the forthcoming results which are new. As we start doing analysis rather than algebra, the assumptions (0.3) and (0.11)-(0.13) or (0.14), (0.15), will start playing a key role.

We begin quoting a lemma which is [10, Theorem 9.7.1].

Lemma 3.1. *Assuming* (0.3)*, we get in the weak-* convergence of measures*

$$(*) - \lim_n \frac{P(.,\alpha_n)}{|\phi_n|^2} dm = d\mu.$$

The two theorems below address the L^2-convergence of Schur remainders under different assumptions. Recall $Acc\,(\alpha_k)$ is the set of accumulation points of (α_k).

Theorem 3.2. *Let* (0.3) *be in force and* $\lim_k |\alpha_k| = 1$*. Assume that* (0.11)-(0.13) *hold. Then*

$$\lim_k \int_{\mathbb{T}} |f_k|^2 P(.,\alpha_k) dm = 0. \tag{3.1}$$

If (α_k) *accumulates nontangentially on* $Acc(\alpha_k) \cap \mathbb{T}$*, then one can replace* (0.11) *and* (0.12) *with* (0.14)*.*

Proof. It is enough to prove (3.1) for any subsequence (α_{n_k}), converging to $\alpha \in Acc\,(\alpha_k)$. For simplicity, the subsequence is still denoted by (α_k).

By Theorem 2.7, we get

$$|\phi_n|^2 \mu' (1 + |f_n|^2 - 2Re(\zeta_n \frac{\phi_n}{\phi_n^*} f_n)) = (1 - |f_n|^2) P(.,\alpha_n)$$

and, consequently,

$$|f_n|^2 = \frac{P(.,\alpha_n) - |\phi_n|^2\mu'}{P(.,\alpha_n) + |\phi_n|^2\mu'} + \frac{2|\phi_n|^2\mu' Re(\zeta_n \frac{\phi_n}{\phi_n^*} f_n)}{P(.,\alpha_n) + |\phi_n|^2\mu'}.$$

Hence, we obtain

$$|f_n|^2 = \frac{P(.,\alpha_n) - |\phi_n|^2\mu'}{P(.,\alpha_n) + |\phi_n|^2\mu'} - \frac{P(.,\alpha_n) - |\phi_n|^2\mu'}{P(.,\alpha_n) + |\phi_n|^2\mu'} Re\left(\zeta_n \frac{\phi_n}{\phi_n^*} f_n\right) + Re\left(\zeta_n \frac{\phi_n}{\phi_n^*} f_n\right).$$

Since $\zeta_n(\alpha_n) = 0$, we get by harmonicity

$$\int_{\mathbb{T}} Re\left(\zeta_n \frac{\phi_n}{\phi_n^*} f_n\right) P(.,\alpha_n) dm = 0,$$

and

$$\int_{\mathbb{T}} |f_n|^2 P(.,\alpha_n) dm = \int_{\mathbb{T}} \frac{P(.,\alpha_n) - |\phi_n|^2\mu'}{P(.,\alpha_n) + |\phi_n|^2\mu'} \left(1 - Re\left(\zeta_n \frac{\phi_n}{\phi_n^*} f_n\right)\right) P(.,\alpha_n) dm.$$

Obviously,

$$\left|1 - Re\left(\zeta_n \frac{\phi_n}{\phi_n^*} f_n\right)\right| \le 2$$

and we get

$$\int_{\mathbb{T}} |f_n|^2 P(.,\alpha_n) dm \le 2 \int_{\mathbb{T}} \left|1 - \frac{2|\phi_n|^2\mu'}{P(.,\alpha_n) + |\phi_n|^2\mu'}\right| P(.,\alpha_n) dm. \tag{3.2}$$

Let

$$g_n = \frac{2|\phi_n|^2\mu'}{P(.,\alpha_n)+|\phi_n|^2\mu'}. \tag{3.3}$$

Using that $4x^2/(1+x)^2 \leq x$ for $x \geq 0$, we deduce

$$\begin{aligned}\int_{\mathbb{T}} g_n^2 P(.,\alpha_n)dm &= \int_{\mathbb{T}} \frac{4(|\phi_n|^2\mu' P(.,\alpha_n)^{-1})^2}{(1+|\phi_n|^2\mu' P(.,\alpha_n)^{-1})^2} P(.,\alpha_n)dm \\ &\leq \int_{\mathbb{T}} |\phi_n|^2\mu' P(.,\alpha_n)^{-1}P(.,\alpha_n)dm \\ &= \int_{\mathbb{T}} |\phi_n|^2\mu' dm \leq \int_{\mathbb{T}} |\phi_n|^2 d\mu = 1.\end{aligned}$$

Therefore, by the Schwarz inequality, it follows that

$$\int_{\mathbb{T}} g_n P(.,\alpha_n)dm \leq \left(\int_{\mathbb{T}} g_n^2 P(.,\alpha_n)dm\right)^{1/2} \leq 1. \tag{3.4}$$

Furthermore, again by the Schwarz inequality,

$$\begin{aligned}\int_{\mathbb{T}} \sqrt{\mu'}P(.,\alpha_n)dm &= \int_{\mathbb{T}} \frac{\sqrt{2}|\phi_n|\sqrt{\mu'}\sqrt{P(.,\alpha_n)}}{\sqrt{P(.,\alpha_n)+|\phi_n|^2\mu'}} \frac{\sqrt{P(.,\alpha_n)+|\phi_n|^2\mu'}\sqrt{P(.,\alpha_n)}}{\sqrt{2}|\phi_n|}dm \\ &\leq \left(\int_{\mathbb{T}} g_n P(.,\alpha_n)dm\right)^{1/2}\left(\frac{1}{2}\int_{\mathbb{T}}\left(\frac{P(.,\alpha_n)}{|\phi_n|^2}+\mu'\right)P(.,\alpha_n)dm\right)^{1/2}.\end{aligned}$$

Recall that the ORFs $(\phi_k)_{0\leq k\leq n}$ are orthonormal in $L^2\left(\frac{P(.,\alpha_n)}{|\phi_n|^2}dm\right)$ and, consequently,

$$\int_{\mathbb{T}} f\,\frac{P(.,\alpha_n)}{|\phi_n|^2}dm = \int_{\mathbb{T}} f\,d\mu$$

for $f \in \mathcal{L}_n + \overline{\mathcal{L}_n}$. Obviously, $P(z,\alpha_n) = z/(z-\alpha_n) + \bar{\alpha_n}z/(1-\bar{\alpha_n}z)$ lies in the latter space and

$$\int_{\mathbb{T}} P(.,\alpha_n)\frac{P(.,\alpha_n)}{|\phi_n|^2}dm = \int_{\mathbb{T}} P(.,\alpha_n)d\mu. \tag{3.5}$$

Using (3.5), we arrive at

$$\int_{\mathbb{T}} \sqrt{\mu'}P(.,\alpha_n)dm \leq \left(\int_{\mathbb{T}} g_n P(.,\alpha_n)dm\right)^{1/2}\left(\int_{\mathbb{T}} P(.,\alpha_n)d\mu\right)^{1/2}. \tag{3.6}$$

Recall now that (α_n) converges to $\alpha \in \mathbb{T}$. By hypothesis, μ' is continuous at α and there is no singular component μ_s in a neighborhood of this point. Thus, passing to the inferior limit in (3.6), we obtain

$$\sqrt{\mu'(\alpha)} \leq \sqrt{\mu'(\alpha)}\liminf_n \left(\int_{\mathbb{T}} g_n P(.,\alpha_n)dm\right)^{1/2}.$$

Moreover, by Fatou's theorem, the same conclusion holds if we assume (0.14) instead of (0.11)-(0.12) provided that (α_k) accumulates nontangentially on $Acc(\alpha_k)\cap\mathbb{T}$. Therefore, since $\mu'(\alpha) > 0$,

$$\liminf_n \int_{\mathbb{T}} g_n P(.,\alpha_n)dm \geq 1.$$

Combining this inequality with (3.4), we see that

$$\lim_n \int_{\mathbb{T}} g_n P(., \alpha_n) dm = \lim_n \int_{\mathbb{T}} g_n^2 P(., \alpha_n) dm = 1, \tag{3.7}$$

and subsequently that

$$\begin{aligned}\lim_n \int_{\mathbb{T}} (1-g_n)^2 P(., \alpha_n) dm &= \int_{\mathbb{T}} P(., \alpha_n) dm - 2 \lim_n \int_{\mathbb{T}} g_n P(., \alpha_n) dm \\ &+ \lim_n \int_{\mathbb{T}} g_n^2 P(., \alpha_n) dm = 0.\end{aligned}$$

With the Schwarz inequality and (3.2), we finish the proof of the first part of the theorem. □

Remark 3.3. *As a partial converse,* (3.1) *implies that* $|f| < 1$ *a.e. on* $Acc(\alpha_k) \cap \mathbb{T}$.

Indeed, observe if $|f| = 1$ a.e. on $E \subset Acc(\alpha_k) \cap \mathbb{T}, |E| > 0$, that $|f_n| = 1$ a.e. on E by Theorem 2.7. The Lebesgue's theorem says that the set of density points of E coincides with E up to a set of Lebesgue measure zero. Comparing the Poisson kernel and the box kernel, we see that the integral in (3.1) cannot go to zero if we pick for α a density point of E.

A similar convergence holds when the (α_n) are compactly included in $\mathbb{D}$. The statement below may seem strange, since the Poisson kernel $P(., \alpha_n)$ is bounded from above and below and therefore superfluous. However, it is convenient to prove the theorem in this form to have it team up with Theorem 3.2 in order to produce Corollary 3.5.

Theorem 3.4. *Let the sequence* (α_k) *be compactly included in* $\mathbb{D}$. *Then,* $|f| < 1$ *a.e. on* $\mathbb{T}$ *if and only if*

$$\lim_n \int_{\mathbb{T}} |f_n|^2 P(., \alpha_n) dm = 0. \tag{3.8}$$

Proof. The "if" part is trivial since $|f_n| = 1$ wherever $|f| = 1$, so we focus on the "only if". As a preliminary, notice that if I is an open arc on $\mathbb{T}$ such that μ has no mass at the end-points of I, it holds that

$$\limsup_n \int_I \frac{P(., \alpha_n)}{|\phi_n|^2} dm \le \mu(I). \tag{3.9}$$

Indeed, in this case, any nested sequence of open arcs I_m decreasing to $\overline{I}$ is such that $\lim_m \mu(I_m) = \mu(\overline{I}) = \mu(I)$. Therefore by the Tietze-Urysohn theorem, there is to each $\varepsilon > 0$ a non-negative function $h_I \in \mathcal{C}(\mathbb{T})$ such that $h_I = 1$ on $\overline{I}$ and $\int_{\mathbb{T}} h_I d\mu \le \mu(I) + \varepsilon$. Obviously

$$\int_I \frac{P(., \alpha_n)}{|\phi_n|^2} dm \le \int_{\mathbb{T}} h_I \frac{P(., \alpha_n)}{|\phi_n|^2} dm,$$

and using Lemma 3.1

$$\lim_n \int_{\mathbb{T}} h_I \frac{P(., \alpha_n)}{|\phi_n|^2} dm = \int_{\mathbb{T}} h_I d\mu \le \mu(I) + \varepsilon.$$

Since ε was arbitrary, this settles the preliminary. Next, define g_n as in (3.3). Arguing as in the previous theorem, we see that equation (3.4) still

holds. Now, it is enough to show that the conclusion of the theorem holds for some infinite subsequence of each sequence of integers. Thus, by Helly's theorem, we are left to establish (3.8) along a subsequence n_k such that $\alpha_{n_k} \to \alpha \in Acc(\alpha_k)$, $\alpha \in \mathbb{D}$, and having the property that g_{n_k} converges to $g \in L^\infty(\mathbb{T})$ in the $*$-weak sense. Clearly $0 \leq g \leq 1$ for the same is true of g_{n_k}. Pick $\xi \in \mathbb{T}$ a Lebesgue point of both g and μ, and let (I_m) be a nested sequence of open arcs decreasing to $\{\xi\}$ such that μ has no mass at the end-points of any I_m. For each m, by the Schwarz inequality,

$$\begin{aligned} \frac{1}{|I_m|}\int_{I_m}\sqrt{\mu'}dm &= \frac{1}{|I_m|}\int_{I_m}\frac{\sqrt{2}|\phi_{n_k}|\sqrt{\mu'}}{\sqrt{P(.,\alpha_{n_k})+|\phi_{n_k}|^2\mu'}}\frac{\sqrt{P(.,\alpha_{n_k})+|\phi_{n_k}|^2\mu'}}{\sqrt{2}|\phi_{n_k}|}dm \\ &\leq \left(\frac{1}{|I_m|}\int_{I_m} g_{n_k}dm\right)^{1/2}\left(\frac{1}{2|I_m|}\int_{I_m}\left(\frac{P(.,\alpha_{n_k})}{|\phi_{n_k}|^2}+\mu'\right)dm\right)^{1/2}. \end{aligned} \tag{3.10}$$

Passing to the limit in (3.10) as $n_k \to \infty$ and using (3.9), we obtain

$$\frac{1}{|I_m|}\int_{I_m}\sqrt{\mu'}dm \leq \left(\frac{1}{|I_m|}\int_{I_m} gdm\right)^{1/2}\left(\frac{1}{2}\frac{\mu(I_m)}{|I_m|}+\frac{1}{2|I_m|}\int_{I_m}\mu'dm\right)^{1/2}.$$

Letting now $m \to \infty$ yields

$$\sqrt{\mu'(\xi)} \leq \sqrt{g(\xi)}\left(\frac{1}{2}\mu'(\xi)+\frac{1}{2}\mu'(\xi)\right)^{1/2} \leq \sqrt{g(\xi)}\sqrt{\mu'(\xi)}. \tag{3.11}$$

By Lebesgue's theorem almost every $\xi \in \mathbb{T}$ satisfies our requirements, and from our assumption that $|f| < 1$ we have $\mu' > 0$, a.e. on $\mathbb{T}$. Consequently $g \geq 1$ by (3.11) hence in fact $g = 1$, a.e. on $\mathbb{T}$. Recalling that $\lim_n P(.,\alpha_n) = P(.,\alpha)$ uniformly on $\mathbb{T}$, we obtain (3.7) from (3.4) and conclude as in Theorem 3.2. □

Corollary 3.5. *Let* (0.3)*,* (0.11)*-*(0.13) *hold and* $|f| < 1$ *a.e. on* $\mathbb{T}$*. Then*

$$\lim_k \int_{\mathbb{T}} |f_k|^2 P(.,\alpha_k)dm = 0.$$

When (α_k) *accumulates nontangentially on* $Acc(\alpha_k)\cap\mathbb{T}$*, assumptions* (0.11)*-*(0.12) *may be replaced by* (0.14)*.*

Proof. It is readily checked that Theorems 3.2 and 3.4 remain valid for subsequences. If the corollary did not hold, it would contradict one of them. □

A closer look at the proof of Theorem 3.2 shows that assumption (0.13) is not really necessary. If $\alpha \in Acc(\alpha_k) \cap \mathbb{T}$ and $\lim_k \alpha_k = \alpha$, all we need is

$$\lim_k \int_{\mathbb{T}} P(.,\alpha_k)\, d\mu_s = 0.$$

For instance if μ_s is a Dirac mass at α and the α_k converge tangentially to α, this could still hold.

4. Convergence of Wall rational functions A_n/B_n

We now discuss different kinds of convergence for the WRFs. This is essentially an interpretation of the results in the previous section, except that we appeal at some point to Proposition 5.6 and Theorem 5.8. The reader will easily convince himself that there is no loophole, *i.e.* that these do not use any result of the present section.

4.1. **Convergence on compact subsets and w.r.t. pseudohyperbolic distance.** Let us begin with an old result that goes back to [46].

Theorem 4.1. *Let* (0.3) *hold. Then A_n/B_n converges to f uniformly on compact subsets of* $\mathbb{D}$.

Proof. As (A_n/B_n) is a family of Schur functions, it is normal. Therefore a subsequence that converges uniformly on compact subsets of $\mathbb{D}$ can be extracted from any subsequence. Let g be the limit of such a subsequence. As $(A_n/B_n)(\alpha_k) = f(\alpha_k)$ for all $\leq n+1$, $f(\alpha_k) = g(\alpha_k)$ for all k. So, the function $f - g \in H^\infty$ vanishes on (α_k) hence it is zero by assumption (0.3). Thus, f is the only limit point. □

Recall that the pseudohyperbolic distance ρ on $\mathbb{D}$ is defined by $\rho(z, w) = |z - w|/|1 - \bar{w}z|$ and it is trivially invariant under Möbius transforms of $\mathbb{D}$.

Theorem 4.2. *Under the assumptions of Corollary 3.5, it holds that*

$$\lim_n \int_{\mathbb{T}} \rho\left(f, \frac{A_n}{B_n}\right)^2 P(., \alpha_{n+1}) dm = 0.$$

Proof. The invariance of the pseudohyperbolic distance under Möbius transforms and relations (1.3) and (1.4) show that

$$\rho\left(f, \frac{A_n}{B_n}\right) = \rho\left(\tau_0 \circ \cdots \circ \tau_n(f_{n+1}), \tau_0 \circ \cdots \circ \tau_n(0)\right) = \rho(f_{n+1}, 0) = |f_{n+1}|.$$

Corollary 3.5 finishes the proof. □

4.2. **Convergence w.r.t. the hyperbolic metric.** In the disk, the hyperbolic metric is defined by

$$\mathfrak{P}(z, \omega) = \log\left(\frac{1 + \rho(z, \omega)}{1 - \rho(z, \omega)}\right). \tag{4.1}$$

Here is an analogue of the "only if" part of Theorem 2.6 from [22].

Theorem 4.3. *Let* (0.3), (0.11)-(0.13) *be in force, and* $\mu \in$ (S). *Then*

$$\lim_n \int_{\mathbb{T}} \mathfrak{P}\left(f, \frac{A_n}{B_n}\right)^2 P(., \alpha_{n+1}) dm = 0.$$

If (α_k) accumulates nontangentially on $Acc(\alpha_k) \cap \mathbb{T}$, then it is enough to assume instead of (0.11) *and* (0.12) *that* (0.15) *holds.*

Proof. We already saw that $\rho(f, A_n/B_n) = |f_{n+1}|$ whence

$$\mathfrak{P}\left(f, \frac{A_n}{B_n}\right) = \log\left(\frac{1 + |f_{n+1}|}{1 - |f_{n+1}|}\right). \tag{4.2}$$

By Theorem 2.7,

$$|\phi_n^*|^2|S|^2\frac{|1-\bar\alpha_n\xi|^2}{1-|\alpha_n|^2}=\frac{1-|f_n|^2}{|1-\zeta_n\frac{\phi_n}{\phi_n^*}f_n|^2}, \tag{4.3}$$

a.e. on $\mathbb{T}$. If g is a Schur function, then $1-g\in H^\infty$ and $\mathrm{Re}\,(1-g)>0$, therefore $1-g$ is an outer function in $H^\infty(\mathbb{D})$ (see [18], Corollary 4.8). Consequently,

$$\int_{\mathbb{T}}\log|1-g|^2P(.,\alpha_n)dm=\log|1-g(\alpha_n)|^2,$$

and, putting $g=\zeta_n\frac{\phi_n}{\phi_n^*}f_n$, we get

$$\int_{\mathbb{T}}\log|1-\zeta_n\frac{\phi_n}{\phi_n^*}f_n|^2P(.,\alpha_n)dm=0.$$

Using the previous equality and (4.3), we see that

$$\int_{\mathbb{T}}\log\left(|\phi_n^*|^2|S|^2\frac{|1-\bar\alpha_n\xi|^2}{1-|\alpha_n|^2}\right)P(\xi,\alpha_n)dm(\xi)=\int_{\mathbb{T}}\log(1-|f_n|^2)P(\xi,\alpha_n)dm(\xi).$$

Since $\log|\phi_n^*|$, $\log|S|$, and $\log|1-\bar\alpha_n\xi|$ are harmonic in $\mathbb{D}$, we continue as

$$\log\big(|\phi_n^*(\alpha_n)|^2|S(\alpha_n)|^2(1-|\alpha_n|^2)\big)=\int_{\mathbb{T}}\log(1-|f_n|^2)P(.,\alpha_n)dm, \tag{4.4}$$

and, by Theorem 5.8 to come, we deduce that

$$\lim_n\int_{\mathbb{T}}\log(1-|f_n|^2)P(.,\alpha_n)dm=0. \tag{4.5}$$

Since $\log(1+x)\le x$ for $x>-1$, we have

$$0\le|f_n|^2\le-\log(1-|f_n|^2),\quad 0\le\log(1+|f_n|)\le|f_n|.$$

Therefore, by the first inequality above and (4.5),

$$\lim_n\int_{\mathbb{T}}|f_n|^2P(.,\alpha_n)dm=0.$$

From this, with the help of the second inequality and the Schwarz inequality,

$$\lim_n\int_{\mathbb{T}}\log(1+|f_n|)P(.,\alpha_n)dm=0.$$

Since $\log(1-|f_n|^2)=\log(1-|f_n|)+\log(1+|f_n|)$, we now see that

$$\lim_n\int_{\mathbb{T}}\log(1-|f_n|)P(.,\alpha_n)dm=0.$$

Referring to (4.2), we finish the proof. □

4.3. **Convergence in $L^2(\mathbb{T})$.** The next theorem follows easily from Corollary 3.5. We begin with

Lemma 4.4. *For $z\in\mathbb{T}$, we have*

$$\left|f(z)-\frac{A_n}{B_n}(z)\right|=|f_{n+1}(z)|\left|1-\frac{A_n}{B_n}(z)\overline{f(z)}\right|\le2|f_{n+1}(z)|.$$

Proof. The equality follows from $\xi_{n+1}f_{n+1}=(A_n-B_nf)/(fA_n^*-B_n^*)$ which is inverse to (1.9). The inequality follows because f and A_n/B_n are Schur. □

Theorem 4.5. *The limiting relation*

$$\lim_n \int_{\mathbb{T}} \left| f - \frac{A_n}{B_n} \right|^p P(., \alpha_{n+1})\, dm = 0 \tag{4.6}$$

holds in the following cases:

1. *if $p \geq 2$ under the assumptions of Corollary 3.5,*
2. *for $1 \leq p < \infty$ if $Acc\,(\alpha_k) \cap \mathbb{T} = \emptyset$ and $|f| < 1$ a.e. on $\mathbb{T}$.*

Proof. This is immediate from Lemma 4.4, Corollary 3.5, the fact that $|f_n| \leq 1$, the existence of pointwise a.e. converging subsequences in L^2-convergent sequences, and the dominated convergence theorem. □

4.4. **Uniform convergence.** When μ is sufficiently smooth, the previous L^p-convergence is uniform.

Theorem 4.6. *Let* (0.3) *hold and $d\mu = \mu' dm$ be absolutely continuous on $\mathbb{T}$. Assume that $\mu' \in W^{1-1/p,p}(\mathbb{T})$ with $p > 4$. If $\mu' > 0$ on some neighborhood $\mathcal{O}(Acc(\alpha_k) \cap \mathbb{T})$, then*

$$\lim_n \left\| \left(f - \frac{A_n}{B_n} \right) \sqrt{P(., \alpha_{n+1})} \right\|_\infty = 0. \tag{4.7}$$

Proof. From (0.4) and (2.15), one easily computes

$$f(z) - A_n/B_n(z) = 2\frac{F_\mu \phi^*_{n+1}(z) - \psi^*_{n+1}(z)}{z(1 + F_\mu(z))(\psi^*_{n+1}(z) + \phi^*_{n+1}(z))}.$$

Besides, we observe from (2.11) that

$$\left|\phi^*_{n+1}(z) + \psi^*_{n+1}(z)\right|^2 = \left|\phi^*_{n+1}(z)\right|^2 + \left|\psi^*_{n+1}(z)\right|^2 + 2P(z, \alpha_{n+1}), \quad z \in \mathbb{T}.$$

From this and the fact that $\mathrm{Re}F_\mu \geq 0$, we obtain on $\mathbb{T}$ that

$$|(f - A_n/B_n)\sqrt{P(., \alpha_{n+1})}| \leq \sqrt{2}\, |F_\mu \phi^*_{n+1} - \psi^*_{n+1}|.$$

Observing that μ' is continuous on $\mathbb{T}$ since $p > 4$, we may apply Corollary 5.14 to the effect that $F_\mu \phi^*_n - \psi^*_n$ converges to zero uniformly on $\mathbb{T}$. □

5. A Szegő-type problem

In this final section, we study the asymptotic behavior of ORFs looking for an analogue of the Szegő theorem in the rational setting when (α_k) may approach $\mathbb{T}$. The results are of a novel type and, we hope, worthy for themselves. We need them also to complete the proofs of Theorems 4.3 and 4.6.

5.1. **Preliminaries.** With π_n defined as in (0.9), we denote by $\mathcal{P}_n\left(d\mu/|\pi_n|^2\right) \subset L^2\left(d\mu/|\pi_n|^2\right)$ the subspace of polynomials of degree at most n. The space $H^2\left(d\mu/|\pi_n|^2\right)$ is the closure of all polynomials in $L^2\left(d\mu/|\pi_n|^2\right)$. The reproducing kernels of $H^2\left(d\mu/|\pi_n|^2\right)$ and $\mathcal{P}_n\left(d\mu/|\pi_n|^2\right)$, are denoted by E_n and R_n, respectively. Since there will be several measures involved, we indicate the dependence in square brackets when necessary. For example, we may write $\phi_n[\mu], E_n[\mu], R_n[\mu]$, or $S[\mu]$, see (0.16). We also put $d\mu_n := d\mu/|\pi_n|^2$.

Proposition 5.1. *Let $\mu \in$ (S) be absolutely continuous. Then,*

$$E_n[\mu](\xi,\omega) = \frac{1}{1-\xi\bar{\omega}} \frac{\pi_n(\xi)\overline{\pi_n(\omega)}}{S(\xi)\overline{S(\omega)}}.$$

The proof is straightforward and stems from the density of polynomials in $H^2\left(d\mu/|\pi_n|^2\right)$, the Cauchy formula, and the identity $|S|^2 = \mu'$ on $\mathbb{T}$.

Proposition 5.2. *The following identity holds:*

$$|\pi_n\phi_n^*| = \frac{|R_n(.,\alpha_n)|}{\|R_n(.,\alpha_n)\|_{L^2(d\mu_n)}}. \tag{5.1}$$

Proof. Let p_{n-1} be a polynomial of degree at most $n-1$. As ϕ_n is orthogonal to $\mathcal{L}_{n-1}$, we have

$$\int_{\mathbb{T}} \overline{\phi_n}\frac{p_{n-1}}{\pi_{n-1}} d\mu = 0.$$

On the other hand, since $\overline{\phi_n} = (\phi_n)_*$ and $1/t = \bar{t}$ on $\mathbb{T}$,

$$\begin{aligned}
\int_{\mathbb{T}} \overline{\phi_n}\frac{p_{n-1}}{\pi_{n-1}} d\mu &= \int_{\mathbb{T}} \phi_n^*(t)\frac{\pi_n(t)}{t^n\overline{\pi_n(t)}}\frac{p_{n-1}(t)(1-\bar{\alpha}_n t)}{\pi_n(t)} d\mu(t) \\
&= \int_{\mathbb{T}} \pi_n(t)\phi_n^*(t)\bar{t}^{n-1}p_{n-1}(t)(\bar{t}-\bar{\alpha}_n)\frac{d\mu(t)}{|\pi_n(t)|^2} \\
&= \int_{\mathbb{T}} \pi_n(t)\phi_n^*(t)\overline{\left(t^{n-1}\overline{p_{n-1}\left(\frac{1}{\bar{t}}\right)}(t-\alpha_n)\right)}\frac{d\mu(t)}{|\pi_n(t)|^2}.
\end{aligned}$$

As $t^{n-1}\overline{p_{n-1}\left(1/\bar{t}\right)}$ ranges over $\mathcal{P}_{n-1}(z)$ when p_{n-1} does, $\pi_n\phi_n^*$ is μ_n-orthogonal to every polynomial of degree $\leq n$ that vanishes at α_n. This is also true of $R_n(.,\alpha_n)$, hence $\pi_n\phi_n^*$ and $R_n(.,\alpha_n)$ are proportional. Since the right-hand side of (5.1) and $\pi_n\phi_n^*$ have unit norm in $L^2(d\mu_n)$, we are done. $\square$

In the following corollary μ is not necessarily absolutely continuous.

Corollary 5.3. *For $\mu \in$ (S) and $n \geq 1$, we have that*

$$|\phi_n^*(\alpha_n)|^2|S(\alpha_n)|^2(1-|\alpha_n|^2) = \frac{R_n(\alpha_n,\alpha_n)}{E_n[\mu_{ac}](\alpha_n,\alpha_n)} \leq 1. \tag{5.2}$$

Proof. By elementary properties of reproducing kernels, we get

$$\|R_n(.,\alpha_n)\|^2_{L^2(d\mu_n)} = R_n(\alpha_n,\alpha_n), \quad \text{and} \quad \|E_n(.,\alpha_n)\|^2_{L^2(d\mu_n)} = E_n(\alpha_n,\alpha_n).$$

Therefore, from Proposition 5.2, we obtain

$$|\pi_n(\alpha_n)\phi_n^*(\alpha_n)|^2 = \frac{|R_n(\alpha_n,\alpha_n)|^2}{\|R_n(.,\alpha_n)\|^2_{L^2(d\mu_n)}} = R_n(\alpha_n,\alpha_n)$$

hence the equality in (5.2) from the formula for $E_n[\mu_{ac}]$ in Proposition 5.1.

Observing now that $||.||_{L^2(d\mu_{ac})} \leq ||.||_{L^2(d\mu)}$, we get a contractive injection

$$H^2\left(d\mu/|\pi_n|^2\right) \subset H^2\left(d\mu_{ac}/|\pi_n|^2\right),$$

from which it follows easily that $E_n[\mu](w,w) \leq E_n[\mu_{ac}](w,w)$, for $w \in \mathbb{D}$.

Since $R_n(.,\alpha_n)$ is the orthogonal projection of $E_n[\mu](.,\alpha_n)$ on $\mathcal{P}_n\left(d\mu/|\pi_n|^2\right)$

$$\|R_n(.,\alpha_n)\|^2_{L^2(d\mu_n)} \leq \|E_n[\mu](.,\alpha_n)\|^2_{L^2(d\mu_n)},$$

and therefore

$$\frac{R_n(\alpha_n,\alpha_n)}{E_n[\mu_{ac}](\alpha_n,\alpha_n)} \leq \frac{R_n(\alpha_n,\alpha_n)}{E_n[\mu](\alpha_n,\alpha_n)} \leq 1,$$

as desired. □

It is a well-known theorem of Beurling that [18, Ch. II, Theorem 7.1] that functions of the form Sp, with p a polynomial, are dense in $H^2(\mathbb{D})$ when S is outer. We shall need a local refinement of this result (compare to [18, Ch. 2, Theorem 7.4]) where, in addition, polynomials get replaced by functions in $\mathcal{L}_n$.

Lemma 5.4. *Let* (0.3) *hold and $\mathcal{O}$ be open in $\mathbb{T}$. Let $S \in H^2(\mathbb{D})$ be an outer function which is continuous on $\mathcal{O}$ with $|S| > \delta > 0$ there. Then, to every compact $K \subset \mathcal{O}$, there is a sequence of rational functions $R_m \in \mathcal{L}_m$ such that*

(i) $\|1 - R_m S\| \to 0$ *as* $m \to \infty$,
(ii) the functions $1 - R_m S$ *go to zero uniformly on* K.

Proof. Recall that $\log|S| \in L^1(\mathbb{T})$ and put $u_n = \min\{a_n, -\log|S|\}$, where $a_n > 0$ tends to $+\infty$ so fast that

$$\sum_{n=0}^{\infty}\left(1 - \exp\left(\int_{\mathbb{T}}(u_n + \log|S|)\,dm\right)\right) < \infty. \tag{5.3}$$

Let S_n be the outer function such that $|S_n| = e^{u_n}$ on $\mathbb{T}$, normalized so that $S_n(0) > 0$. Then $S_n \in H^\infty(\mathbb{D})$ and $|S_nS| \leq 1$ on $\mathbb{T}$ with $|S_nS| = 1$ on $\mathcal{O}$ for n large enough, therefore we can write

$$SS_n(z) = \exp\left(\int_{\mathbb{T}\setminus\mathcal{O}} \frac{t+z}{t-z}\log|SS_n|\,dm(t)\right), \tag{5.4}$$

showing that SS_n extends analytically across $\mathcal{O}$ to an analytic function G_n on $\overline{\mathbb{C}}\setminus(\mathbb{T}\setminus\mathcal{O})$. Moreover, (G_n) is a normal family since $|\log|S_n|| \leq |\log|S||$ on $\mathbb{T}$. Besides, expanding $\|1 - S_nS\|^2$ and using (5.3), we obtain

$$\sum_{n=0}^{\infty}\|1 - S_nS\|^2 \leq 2\sum_{n=0}^{\infty}(1 - S_n(0)S(0)) < \infty \tag{5.5}$$

so that, by the Borel-Cantelli lemma, SS_n converges to 1 a.e. on $\mathbb{T}$. Then, by normality, SS_n converges to 1 locally uniformly on $\mathcal{O}$.

Next, fix a compact $K \subset \mathcal{O}$ and let $S_{n,r}(z) := S_n(rz)$ for $0 < r < 1$. As $S_{n,r} = P_{rz} * S_n$, and since $S_n \in L^\infty(\mathbb{T})$ is continuous on $\mathcal{O}$ where it equals G_n/S, it follows from standard properties of Poisson integrals [18, Ch. 2] that $S_{n,r}$ converges to S_n boundedly pointwise a.e. on $\mathbb{T}$ and locally uniformly on $\mathcal{O}$ as $r \to 1$. In particular, $S_{n,r}S$ converges to S_nS in $L^2(\mathbb{T})$ for fixed n as $r \to 1$. Hence to each n there is r_n such that, say,

$$\begin{cases} \|1 - S_{n,r_n}S\| & < \quad \|1 - S_nS\| + 2^{-n}, \\ \sup_K |S_{n,r_n} - S_n| & < \quad 1/n. \end{cases}$$

Clearly S_{n,r_n} lies in $A(\mathbb{D})$, therefore is can be uniformly approximated on $\mathbb{T}$ by functions from $\cup_k \mathcal{L}_k$ since (0.3) holds. Therefore, to each n, there is an integer m_n and $R_{m_n} \in \mathcal{L}_{m_n}$ such that

$$\left\{ \begin{array}{lcl} \|1 - R_{m_n} S\| & < & \|1 - S_n S\| + 2^{-n}, \\ \sup_K |R_{m_n} - S_n| & < & 1/n. \end{array} \right. \tag{5.6}$$

Without loss of generality, we assume that m_n strictly increases with n. Now, by (5.5) and since $|SS_n| \in L^\infty(\mathbb{T})$, the first relation in (5.6) implies

$$\sum_{n=0}^{\infty} \|1 - R_{m_n} S\|^2 < \infty$$

whence $R_{m_n} S$ converges to 1 in $H^2(\mathbb{D})$ as $n \to \infty$. In another connection,

$$|1 - R_{m_n} S| \leq |1 - S_n S| + |R_{m_n} - S_n||S|$$

and the second relation in (5.6) yields that $R_{m_n} S$ converges uniformly to 1 on K when $m_n \to \infty$. To complete the proof, it remains to put $R_m = R_{m_k}$ where k is the greatest integer such that $m_k \leq m$. □

5.2. **An a priori bound on ORFs.** We derive in this subsection *a priori* estimates for ORFs akin to the classical bounds for orthogonal polynomials [20, Ch. 4, Theorems 4.6, 4.8], [44, Ch. 12, Theorem 12.1.3]. These in fact are new even in the classical polynomial case, as they yield information in cases where μ' vanishes, thereby generalizing some of the results from [40]. Their proof rely on basic properties of the Sobolev spaces $W^{1,p}(\Omega)$. For $1 < p < \infty$ and $\Omega \subset \mathbb{C}$ an open set with boundary $\partial\Omega$, recall that

$$W^{1,p}(\Omega) = \{f \in L^p(\Omega) : ||f||_{L^p(\Omega)} + ||f'||_{L^p(\Omega)} < \infty\},$$

where the derivatives are understood in the distributional sense. If $\partial\Omega$ is piecewise smooth and $\mathcal{D}$ indicates the space of $\mathcal{C}^\infty$ functions with compact support in $\mathbb{C}$, then the restriction $\mathcal{D}|_\Omega$ is dense in $W^{1,p}(\Omega)$. For $g \in W^{1,p}(\Omega)$ and (g_n) a sequence in $\mathcal{D}|_\Omega$ converging to g, one can show that the trace of g_n on $\partial\Omega$ converges in $W^{1-1/p,p}(\Omega)$, see (0.19). This allows one to define the *trace* of $g \in W^{1,p}(\Omega)$ on $\partial\Omega$ as a member of $W^{1-1/p,p}(\partial\Omega)$. With this definition, Stokes' formula holds for Sobolev differential forms just like it does for smooth ones.

We put $\eta = x + iy$ and use the standard notation

$$\frac{\partial}{\partial\eta} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right), \qquad \frac{\partial}{\partial\bar\eta} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

The usual rules of differentiation apply to $\partial/\partial\eta$, $\partial/\partial\bar\eta$, and the relation $\partial V/\partial\bar\eta = 0$ means that V is analytic. We need a function-theoretic lemma.

Lemma 5.5. *Let $I \subset \mathbb{T}$ be an open arc and $\Omega \subset \mathbb{D}$ an open set such that $\overline{\Omega} \cap \mathbb{T} \subset I$. If $g \in H^1(\mathbb{D})$ is such that $g|_I \in W^{1-1/p,p}(I)$ for some $1 < p < \infty$, then $g|_\Omega \in W^{1,p}(\Omega)$.*

Proof. The restriction $g|_I$ extends to a function $h \in W^{1-1/p,p}(\mathbb{T})$ [1, 7.69]. By standard elliptic regularity, there is a harmonic function $U \in W^{1,p}(\mathbb{D})$ such that $U|_\mathbb{T} = h$, where the trace is understood in the Sobolev sense [12]. Any harmonic conjugate V of U in turn belongs to $W^{1,p}(\mathbb{D})$ for $\partial V/\partial\bar\eta = i\partial U/\partial\bar\eta$ by the Cauchy-Riemann equations. Hence the analytic function

$F = U + iV$ lies in $W^{1,p}(\mathbb{D})$, and it follows from the definition that $F(r\eta) \to F(\eta)$ in $W^{1,p}(\mathbb{D})$ as $r \to 1^-$. Thus by the trace theorem, the restriction of F to every circle $\mathbb{T}_r$ centered at 0 of radius $r < 1$ has $W^{1-1/p,p}(\mathbb{T}_r)$ norm at most $C\|F\|_{W^{1,p}(\mathbb{D})}$ for some constant C [1, 7.39]. *A fortiori* then, $F \in H^p(\mathbb{D})$ and its Sobolev trace on $\mathbb{T}$ must coincide with its nontangential limit. Consequently $g - F$ is a $H^1(\mathbb{D})$-function which is pure imaginary on I, therefore it extends analytically by reflection across I. In particular $g - F$ is smooth on a neighborhood of $\overline{\Omega}$ and $g = F + (g - F)$ lies in $W^{1,p}(\Omega)$. □

Our *a priori* bound will depend on the connection between ϕ_n, ψ_n and u_n defined in (2.9). By Proposition 2.4, $zu_n(z) = F_\mu(z)(\phi_n(z))_* - (\psi_n(z))_*$ for $z \in \mathbb{T}$. Multiplying by ϕ_n and taking real parts, we get for $z \in \mathbb{T}$

$$\mu'(z)|\phi_n(z)|^2 = \mathrm{Re}((\psi_n)_*(z)\phi_n(z)) + \mathrm{Re}(z\phi_n(z)u_n(z)),$$

and a short computation using (2.11) gives us

$$\left|\mu'(z)\phi_n(z) - \frac{z\bar{u}_n(z)}{2}\right|^2 = \mu'(z)P(z, \alpha_n) + \left|\frac{u_n(z)}{2}\right|^2.$$

Thus, either $|\mu'(z)\phi_n(z)| \le |u_n(z)|$ or $|\mu'(z)\phi_n(z)|^2/4 < \mu'(z)P(z, \alpha_n) + |u_n(z)|^2/4$. Therefore, for $z \in \mathbb{T}$,

$$\mu'^2(z)|\phi_n(z)|^2 \le |u_n(z)|^2 + \mu'(z)P(z, \alpha_n). \tag{5.7}$$

Proposition 5.6. *Let $\mu \in$ (S) and $I \subset \mathbb{T}$ be an open arc disjoint from* $\operatorname{supp}\mu_s$. *Assume that $S \in W^{1-1/p,p}(I)$ with $4/3 < p < \infty$. Then, to each compact $K \subset I$,*

i) there is a neighborhood $\mathcal{O}(K)$ in I such that $u_n|_{\mathcal{O}(K)} \in W^{1-1/\gamma,\gamma}(\mathcal{O})$ for $1 < \gamma < 4p/(p+4)$, with norm depending on μ, K, and γ only.

ii) If moreover $p > 4$, then $|u_n| \le C$ on $\mathcal{O}(K)$ and $u_n \in H_s(\mathcal{O}(K))$ for $0 < s < (p-4)/2p$, where C and the Hölder constant depend on μ, K, and s only; in particular, from (5.7)*, we obtain for $\xi \in K$*

$$\mu'^2(\xi)|\phi_n(\xi)|^2 \le C + \mu' P(\xi, \alpha_n). \tag{5.8}$$

Proof. We may assume that $K \ne \mathbb{T}$, otherwise the conclusion follows upon writing $\mathbb{T} = K_1 \cup K_2$. Let $J = (e^{i\theta_1}, e^{i\theta_2})$ be an open arc compactly included in I and containing K. Fix $0 < \varepsilon < 1$ and consider the radial segments $c_1 := [e^{i\theta_1}, (1+\varepsilon)e^{i\theta_1}]$, $c_2 := [(1+\varepsilon)e^{i\theta_2}, e^{i\theta_2}]$, and the circular arc $c_3 = \{(1+\varepsilon)e^{i\theta} : \theta_1 \le \theta \le \theta_2\}$. Let $\mathcal{C} = c_1 \cup c_3 \cup c_2$ be the open contour joining $e^{i\theta_1}$ to $e^{i\theta_2}$. Orient the Jordan curve $\Gamma = \mathcal{C} \cup J$ counterclockwise, and let Ω_1 denote its interior. Put $\Omega = \{z \in \mathbb{D};\, 1/\bar{z} \in \Omega_1\}$ for the reflected set. Lemma 5.5 implies that $S \in W^{1,p}(\Omega)$, hence $G(z) := S(1/\bar{z})$ and $H(z) := \overline{S(1/\bar{z})}$ belong to $W^{1,p}(\Omega_1)$. By a classical estimate [16, Theorem 5.4], $H^2(\mathbb{D})$ embeds continuously in $L^\beta(\mathbb{D})$ for $2 < \beta < 4$. *A fortiori* then, $S \in L^\beta(\Omega)$ whence $G, H \in L^\beta(\Omega_1)$ by reflection. In particular, from the Leibnitz rule and Hölder's inequality, it follows since $p > 4/3$ that $GH(z) = |S(1/\bar{z})|^2$ lies in $W^{1,\alpha}(\Omega_1)$ for some $\alpha > 1$. Pick $z \notin \overline{\Omega_1}$ and apply Stokes' theorem on Γ to the differential form $GH(\eta)(\phi_n)_*(\eta)/(\eta(\eta - z))\, d\eta$:

$$\int_{\mathcal{C}\cup J} \frac{GH(\eta)(\phi_n)_*(\eta)}{\eta} \frac{d\eta}{\eta - z} = -\int_{\Omega_1} \frac{(\partial G/\partial\bar{\eta})(\eta)}{\eta - z} \frac{H(\phi_n)_*(\eta)}{\eta}\, d\eta \wedge d\bar{\eta}, \tag{5.9}$$

where we took into account that $H(\phi_n)_*(\eta)/(\eta(\eta-z))$ is analytic on Ω_1, since H and $(\phi_n)_*$ are analytic on $\overline{\mathbb{C}}\setminus\overline{\mathbb{D}}$ while $0,z\notin\overline{\Omega_1}$.

As $GH(\xi)d\xi/\xi = id\mu(\xi)$ on $\bar{J}$ because $\operatorname{supp}\mu_s\cap I=\emptyset$ by assumption, we deduce from (5.9) and (2.9) that, for $z\in\mathbb{D}$

$$
\begin{aligned}
u_n(z) &= 2\int_{\mathbb{T}\setminus J}(\phi_n)_*(\xi)\,\frac{d\mu(\xi)}{\xi-z} - 2i\int_{\mathcal{C}}\frac{GH(\xi)(\phi_n)_*(\xi)}{\xi}\,\frac{d\xi}{\xi-z}\\
&- 2i\int_{\Omega_1}\frac{(\partial G/\partial\bar\eta)(\eta)}{\eta-z}\,\frac{H(\phi_n)_*(\eta)}{\eta}\,d\eta\wedge d\bar\eta.
\end{aligned}
\tag{5.10}
$$

On any $\mathcal{O}(K)$ where z remains at strictly positive distance from $\mathbb{T}\setminus J$, the first integral in the right-hand side of (5.10) is uniformly bounded and smooth by the Schwarz inequality because $||(\phi_n)_*||_{L^2(d\mu)}=1$ (recall that $|(\phi_n)_*|=|\phi_n|$ on $\mathbb{T}$).

Next, since $\phi_nS\in H^2(\mathbb{D})$ and $||\phi_nS||=||\phi_n||_{L^2(d\mu_{ac})}\le 1$, it follows from the Fejèr-Riesz inequality [16, Theorem 3.13] that the L^2-norm of ϕ_nS over any diameter of $\mathbb{D}$ is at most $1/\sqrt{2}$. Also, the L^2-norm of ϕ_nS over the circle centered at zero of radius $1/(1+\varepsilon)$ is less than 1. Hence, by reflection across $\mathbb{T}$, the L^2-norm of $H(\phi_n)_*$ on $\mathcal{C}$ is uniformly bounded. Moreover, since $G\in W^{1,p}(\Omega_1)$, the trace theorem implies that its restriction to $\mathcal{C}$ lies in $W^{1-1/p,p}(\mathcal{C})$. By the embedding theorem for Besov spaces [1, Theorem 7.34], this restriction belongs[1] to $L^{p/(2-p)}(\mathcal{C})$ if $p<2$, to each $L^q(\mathcal{C})$ with $1\le q<\infty$ if $p=2$, and it is bounded if $p>2$. Thus from Hölder's inequality, it follows since $p>4/3$ that the $L^2(\mathcal{C})$-norm of $G|_{\mathcal{C}}$ it at most $C\|G\|_{W^{1,p}(\Omega_1)}$, where C is a constant depending only of p and $\mathcal{C}$. Consequently, by the Cauchy-Schwarz inequality, the second integral in the right-hand side of (5.10) is uniformly bounded and smooth on any $\mathcal{O}(K)$ remaining at positive distance from $\mathcal{C}$.

We turn to the third integral, which is taken with respect to two-dimensional Lebesgue measure since $d\eta\wedge d\bar\eta=-2idxdy$. For $z\in\Omega_1$, observe that the function

$$V(z)=\int_{\Omega_1}\frac{v(\eta)}{\eta-z}d\eta\wedge d\bar\eta$$

lies in $W^{1,\gamma}(\Omega_1)$ whenever $v\in L^\gamma(\Omega_1)$ with $1<\gamma<\infty$. Indeed, if d is the diameter of Ω_1, it holds that $\|V\|_{L^\gamma(\Omega_1)}\le 6d\|v\|_{L^\gamma(\Omega_1)}$ [6, Theorem 4.3.12] while the distributional derivatives $\partial V/\partial\bar z$ and $\partial V/\partial z$ equal respectively to v and the restriction to Ω_1 of $\mathcal{S}\check v$, where $\check v$ is the extension of v by 0 to the whole of $\mathbb{C}$ and $\mathcal{S}$ indicates the Beurling transform [6, Theorem 4.3.10]. Since the latter is a bounded operator on $L^\gamma(\mathbb{C})$ [6, Theorem 4.5.3], it follows that $V\in W^{1,\gamma}(\Omega_1)$ with norm depending on Ω_1 and $\|v\|_{L^\gamma(\Omega_1)}$. Apply this to

$$v(\eta)=\frac{H(\eta)(\phi_n)_*(\eta)}{\eta}\,\frac{\partial G}{\partial\bar\eta}(\eta)$$

so that V becomes the third integral in (5.10), up to the factor $-2i$. On the one hand, $G\in W^{1,p}(\Omega_1)$ so that $\partial G/\partial\bar\eta\in L^p(\Omega_1)$. On the other hand, we pointed out already that $H^2(\mathbb{D})$ embeds in $L^\beta(\mathbb{D})$ for $2<\beta<4$, hence the $L^\beta(\Omega)$-norm of ϕ_nS is uniformly bounded and so is the $L^\beta(\Omega_1)$-norm

[1] It belongs in fact to the Lorentz space $L^{p/(2-p),p}(\mathcal{C})$ that we did note introduce; the latter is included in $L^{p/(2-p)}(\mathcal{C})$ because $p<p/(2-p)$ since $p>1$, see [47, Lemma 1.8.13].

of $H(\phi_n)_*$ by reflection. Therefore by Hölder's inequality, $v \in L^\gamma(\Omega_1)$ with $1/\gamma = 1/p + 1/\beta$. As $p > 4/3$, this allows us to pick γ arbitrarily in the range $(1, 4p/(p+4))$, and assertion $i)$ now follows from the trace theorem.

If $p > 4$ we can pick $2 < \gamma < 4p/(p+4)$, so assertion $ii)$ is a consequence of (5.7) and the fact that $W^{1-1/\gamma,\gamma}(I)$ embeds continuously in $H_{1-2/\gamma}(I)$. $\square$

The importance of the above proposition lies with the fact that the bounds are independent of n and (α_k), except for the presence of $P(., \alpha_n)$ in (5.8) (which reduces to 1 in the classical case).

It is useful to know conditions on $\mu' = |S|^2$ for Proposition 5.6 to apply. Here is a simple criterion.

Lemma 5.7. *Let $1 < p < \infty$ and $\mu \in$ (S). For $I \subset \mathbb{T}$ an open arc, if $\mu'|_I \in W^{1-1/p,p}(I)$ and $0 < \delta \leq \mu'(t) \leq M < \infty$ on I, it holds that $S|_J \in W^{1-1/p,p}(J)$ for each relatively compact subarc $J \subset I$.*

Proof. Let $\varphi \in W^{1-1/p,p}(\mathbb{T})$ coincide with $\mu'/2$ on I with $0 < \delta' \leq \varphi \leq M' < \infty$; such an extension is easily constructed by reflexion across the endpoints of I, see [21, Theorem 1.5.2.3]. As $\varphi \geq \delta' > 0$, we get that $\log \varphi \in W^{1-1/p,p}(I)$ for log is Lipschitz continuous on the range of φ. Therefore $H = \log \varphi + i(\widetilde{\log \varphi}) \in W^{1-1/p,p}(\mathbb{T})$, and so does $S_1 = \exp H$ because exp is Lipschitz continuous on the range of H since $\varphi \leq M'$. Now, S_1 is an outer function having the same modulus as S on I, therefore we see as in (5.4) that S/S_1 extends analytically across I. This entails that $S \in W^{1-1/p,p}(J)$ whenever J is relatively compact in I. $\square$

It is straightforward to check that the product of bounded $W^{1-1/p,p}(I)$-functions again lies in $W^{1-1/p,p}(I)$. Thus if μ' satisfies the conditions of Lemma 5.7, then the conclusion still holds for $\mu_1'(t) = \mu'(t)\Pi_{j=1}^N |t - t_j|^{\lambda_j}$ where $t_1, \ldots t_N \in I$ are distinct, and either $\lambda_j \geq 2$ or $\lambda_j > 2(p-1)/p$ for all j, because the (normalized) outer function with modulus $|t - t_j|^{\lambda_j/2}$ is just $(t - t_j)^{\lambda_j/2}$, where the branch of the power $\lambda_j/2$ is positive for positive arguments. For instance $d\mu(\xi) = |1 - \xi|^2 dm(\xi)$ provides us with an example for which (5.8) holds uniformly on $\mathbb{T}$ although $\mu'(1) = 0$.

5.3. **Convergence of ORFs for Szegő measures.** In this subsection, we assume as always that μ is a finite and positive measure with infinite support on $\mathbb{T}$, but we no longer require it has unit mass. The ORFs and the associated Carathéodory and Szegő functions are defined as before. Because multiplying μ by $\lambda > 0$ results in the multiplication of ϕ_n by $\lambda^{-1/2}$ and of S by $\lambda^{1/2}$, the results below are invariant under such scalings.

Observe that, similarly to the classical situation, the ORF ϕ_n solves the extremal problem

$$\max_{\xi_n \in \mathcal{L}_n,\ ||\xi_n||_\mu \leq 1} \left\{ |a_{n,n}| : \ \xi_n = a_{n,n}\mathcal{B}_n + a_{n,n-1}\mathcal{B}_{n-1} + \cdots + a_{n,0}\mathcal{B}_0 \right\}. \tag{5.11}$$

We denote the value of the problem by $\kappa_n = \kappa_n[\mu]$, *i.e.* $\kappa_n = |\phi_n^*(\alpha_n)|$.

The extremal property (5.11) can also be recast as

$$\kappa_n^{-1} = \min_{\xi_n \in \mathcal{L}_n,\ \xi_n(\alpha_n) = 1} \|\xi_n\|_\mu, \tag{5.12}$$

where the extremal value is uniquely attained at $\xi_n = \phi_n^*/\phi_n^*(\alpha_n)$.

The Szegő-type theorem we shall prove deals with the asymptotic behavior of κ_n as $n \to +\infty$, which entails further asymptotics for ϕ_n^*.

The statement is as follows.

Theorem 5.8. *Let* (0.3), (0.11)-(0.13) *be in force, and* $\mu \in$ (S). *Then*

$$\lim_n |\phi_n^*(\alpha_n)|^2 |S(\alpha_n)|^2 (1 - |\alpha_n|^2) = 1. \tag{5.13}$$

If (α_k) *accumulates nontangentially on* $Acc(\alpha_k) \cap \mathbb{T}$, *then one can replace relations* (0.11) *and* (0.12) *with* (0.15).

The proof of Theorem 5.8 requires several steps. We first look at smooth measures.

Proposition 5.9. *Assume* (0.3) *holds and* $\mu \in$ (S) *is absolutely continuous. If there is an open neighborhood* $\mathcal{O}$ *of* $Acc(\alpha_k) \cap \mathbb{T}$ *where* $\mu' \geq \delta > 0$ *with* $\mu' \in W^{1-1/p,p}(I)$ *for each component* I *of* $\mathcal{O}$, $p > 2$, *then* (5.13) *is valid.*

Proof. Since $R_n(., \alpha_n)$ is the orthogonal projection of $E_n(., \alpha_n)$ on $\mathcal{P}_n(d\mu_n)$, $R_n(., \alpha_n)$ is a polynomial of degree at most n and the minimum

$$\min_{p_n \in \mathcal{P}_n} \|E_n(., \alpha_n) - p_n\|_{L^2(d\mu_n)}$$

is attained exactly for $p_n = R_n(., \alpha_n)$. But

$$\|E_n(., \alpha_n) - p_n\|^2_{L^2(d\mu_n)} = \int_{\mathbb{T}} \left| \frac{1}{1 - \overline{\alpha_n} t} \frac{\overline{\pi_n(\alpha_n)}}{\overline{S(\alpha_n)}} - \frac{p_n(t) S(t)}{\pi_n(t)} \right|^2 dm(t).$$

Hence, the polynomial P_n minimizing

$$\min_{p_n \in \mathcal{P}_n} \left\| \frac{1}{1 - \overline{\alpha_n} t} - \frac{p_n(t) S(t)}{\pi_n(t)} \right\| \tag{5.14}$$

provides us with $R_n(., \alpha_n)$ through the relation

$$R_n(., \alpha_n) = \frac{\overline{\pi_n(\alpha_n)}}{\overline{S(\alpha_n)}} P_n.$$

In view of (5.2), we write

$$|\phi_n^*(\alpha_n)|^2 |S(\alpha_n)|^2 (1 - |\alpha_n|^2) = \left| \frac{P_n(\alpha_n) S(\alpha_n)}{\pi_{n-1}(\alpha_n)} \right|. \tag{5.15}$$

We also have for every polynomial p_n

$$\begin{aligned} \left\| \frac{1}{1 - \bar{\alpha}_n t} - \frac{p_n(t) S(t)}{\pi_n(t)} \right\|^2 &= \left\| \left(1 - \frac{p_n(t) S(t)}{\pi_{n-1}(t)} \right) \frac{1}{t - \alpha_n} \right\|^2 \\ = & \left\| \left(1 - \frac{p_n(\alpha_n) S(\alpha_n)}{\pi_{n-1}(\alpha_n)} \right) \frac{1}{t - \alpha_n} + \left(\frac{p_n(\alpha_n) S(\alpha_n)}{\pi_{n-1}(\alpha_n)} - \frac{p_n(t) S(t)}{\pi_{n-1}(t)} \right) \frac{1}{t - \alpha_n} \right\|^2. \end{aligned}$$

Consequently,

$$\begin{aligned} \left\| \frac{1}{1 - \bar{\alpha}_n t} - \frac{p_n(t) S(t)}{\pi_n(t)} \right\|^2 &= \left| 1 - \frac{p_n(\alpha_n) S(\alpha_n)}{\pi_{n-1}(\alpha_n)} \right|^2 \frac{1}{1 - |\alpha_n|^2} \\ &+ \left\| \left(\frac{p_n(\alpha_n) S(\alpha_n)}{\pi_{n-1}(\alpha_n)} - \frac{p_n(t) S(t)}{\pi_{n-1}(t)} \right) \frac{1}{t - \alpha_n} \right\|^2. \end{aligned} \tag{5.16}$$

Thus, if there is a sequence of polynomials (p_n) satisfying

$$\left\|\frac{1}{1-\bar{\alpha}_n t}-\frac{p_n(t)S(t)}{\pi_n(t)}\right\|^2=o\left(\frac{1}{1-|\alpha_n|^2}\right), \tag{5.17}$$

then we also have (see (5.14))

$$\left\|\frac{1}{1-\bar{\alpha}_n t}-\frac{P_n(t)S(t)}{\pi_n(t)}\right\|^2=o\left(\frac{1}{1-|\alpha_n|^2}\right),$$

and by (5.16)

$$\lim_n \frac{P_n(\alpha_n)S(\alpha_n)}{\pi_{n-1}(\alpha_n)}=1.$$

In this case relation (5.15) gives us the desired limit (5.13).

Note that μ' is bounded on each component I of $\mathcal{O}$, since $W^{1-1/p,p}(I)$ consists of continuous functions for $p>2$. Hence S is continuous on $\mathcal{O}$ by Lemma 5.7, and meets the assumptions of Lemma 5.4. Let K be a compact neighborhood of $Acc(\alpha_k)$ included in $\mathcal{O}$ and $R_n\in\mathcal{L}_n$ be the sequence of rational functions given by the lemma. Put $R_n=p_n/\pi_n$. As $\|1/(1-\bar{\alpha_n}t)\|^2=1/(1-|\alpha_n|^2)$, we get

$$\begin{aligned}\left\|\frac{1}{1-\bar{\alpha}_n t}-\frac{p_{n-1}(t)S(t)}{\pi_n(t)}\right\|^2 \quad &\leq \quad \sup_{t\in\mathbb{T}\setminus K}\frac{1}{|1-\overline{\alpha}_n t|^2}\quad\left\|1-\frac{p_{n-1}S}{\pi_{n-1}}\right\|^2\\ &+\quad\frac{1}{1-|\alpha_n|^2}\quad\left\|1-\frac{p_{n-1}S}{\pi_{n-1}}\right\|^2_{L^\infty(K)}.\end{aligned}$$

Since K is a neighborhood of $Acc(\alpha_k)$, the above supremum is bounded and the first summand in the right-hand side of the equation goes to zero as $n\to\infty$ by the properties of R_n. As to the second summand, it is $o(1/(1-|\alpha_n|^2))$ since $R_{n-1}S$ converges to 1 uniformly on K. Therefore the sequence (p_{n-1}) satisfies (5.17) whence (5.13) holds. □

The assumption on μ' was only to ensure that S is continuous on $\mathcal{O}$. If this is known to be the case, it is not needed.

The next proposition will sharpen Proposition 5.9 in that it allows μ to have a singular part. We need a preparatory lemma.

Lemma 5.10. *Assume* (0.3) *holds and let $E\subset\mathbb{T}$, $|E|=0$, have an open neighborhood $\mathcal{U}$ in $\mathbb{C}$ such that $\overline{\mathcal{U}}\cap Acc(\alpha_k)=\emptyset$. Then, to every $\varepsilon>0$, there exists an integer n_0 and $R_{n_0}\in\mathcal{L}_{n_0}$ such that*

(i) $|R_{n_0}|\leq 2+\varepsilon$ *on* $\mathbb{T}$,
(ii) $|1-R_{n_0}|\leq\varepsilon$ *on* $\mathbb{T}\setminus\mathcal{U}$,
(iii) $|R_{n_0}|\leq\varepsilon$ *on* E,

Proof. Since $E=\operatorname{supp}\mu_s$ has Lebesgue measure zero, one can find $g\in A(\mathbb{D})$ such that $g=1$ on E and $|g|<1$ on $\overline{\mathbb{D}}\setminus E$, *cf.* [18, Ch. 3, Exercise 2]. Pick m so large that $|g^m|<\varepsilon/2$ on $\overline{\mathbb{D}}\setminus\mathcal{U}$. Since (0.3) holds and $(1-g^m)\in A(\mathbb{D})$, we can find n_0 and $R_{n_0}\in\mathcal{L}_{n_0}$ such that $|1-g^m-R_{n_0}|<\varepsilon/2$ on $\overline{\mathbb{D}}$. □

We also take note of the identity

$$F_\mu=S[\mu]/S[\widetilde{\mu}], \tag{5.18}$$

where $\widetilde{\mu}$ is the Herglotz measure of $1/F_\mu$, see (0.4). Indeed, since Carathéodory functions are outer, both sides of (5.18) are outer functions, positive at 0, with equal modulus a.e. on $\mathbb{T}$ as can be readily computed from (0.5).

Proposition 5.11. *Assumptions being as in Proposition 5.9, except that μ may now have a singular part satisfying* (0.13)*, we have that* (5.13) *holds.*

Proof. In view of Corollary 5.3, all we have to prove is that

$$\liminf_n (1-|\alpha_n|^2)|S(\alpha_n)|^2\kappa_n^2[\mu] \geq 1. \tag{5.19}$$

Assume first that $\mu' \geq \delta' > 0$ on an open set $\mathcal{V} \supset \operatorname{supp}\mu_s$ in $\mathbb{T}$ and that the restriction of μ' to each component I of $\mathcal{V}$ lies in $W^{1-1/q,q}(I)$ for some $q > 4$. Then S is continuous on $\mathcal{V}$. Thanks to (0.13), we may require in addition that $\overline{\mathcal{V}} \cap \mathcal{O} = \emptyset$ and also, by the compactness of $\operatorname{supp}\mu_s$, that $\mathcal{V}$ has finitely many components. Fix a neighborhood $\mathcal{W}$ of $\operatorname{supp}\mu_s$ in $\mathbb{T}$ with $\overline{\mathcal{W}} \subset \mathcal{V}$. We can apply Proposition 5.6 to $d\mu_{ac}$ on each component I of $\mathcal{V}$ with $K = \overline{\mathcal{W}} \cap I$ and deduce from (5.8), since $\overline{\mathcal{W}}$ remains at positive distance from (α_k), that each ORF θ_j associated with μ_{ac} and with *any* subsequence (β_l) of (α_k) is bounded by a constant C_0 on $\overline{\mathcal{W}}$, where C_0 is *independent* of j and of the subsequence. In particular, since $\mu' \in L^1(\mathbb{T})$ and $|\theta_j^*| = |\theta_j|$ on $\mathbb{T}$, to any $\varepsilon > 0$ there is $\eta > 0$ such that for $j \in \mathbb{N}$, $(\beta_l) \subset (\alpha_k)$,

$$\int_{\mathcal{W}_1} |\theta_j^*|^2 \mu' \, dm < \varepsilon, \tag{5.20}$$

as soon as $\mathcal{W}_1 \subset \overline{\mathcal{W}}$ has Lebesgue measure less than η.

Pick $\varepsilon > 0$ and let $\mathcal{W}_1 \subset \mathcal{W}$ be an open neighborhood of $\operatorname{supp}\mu_s$ in $\mathbb{T}$ such that $|\mathcal{W}_1| < \eta$. This is possible since $|\operatorname{supp}\mu_s| = 0$. Write $\mathcal{W}_1 = \mathcal{U} \cap \mathbb{T}$ where $\mathcal{U}$ is open in $\mathbb{C}$ and $\overline{\mathcal{U}} \cap (\alpha_k) = \emptyset$. This can be ensured because $\overline{\mathcal{W}} \cap Acc(\alpha_k) = \emptyset$. Apply Lemma 5.10 with $E = \operatorname{supp}\mu_s$, and let $R_{n_0} \in \mathcal{L}_{n_0}$ be as in the lemma. Consider the sequence (θ_j) of ORFs associated with $d\mu_{ac}$ for the truncated sequence $\beta_l = \alpha_{l+n_0}$, $l \geq 1$. Hence for $n > n_0$, we have that $|\theta_{n-n_0}^*(\alpha_n)| = \kappa'_{n-n_0}[\mu_{ac}]$ where the prime in $\kappa'_{n-n_0}[\mu_{ac}]$ indicates that we work with the truncated sequence $(\alpha_k)_{k>n_0}$. By (5.12), we get

$$\kappa_n^{-2}[\mu] \leq \int_{\mathbb{T}} \left| \frac{\theta_{n-n_0}^* R_{n_0}}{\theta_{n-n_0}^*(\alpha_n) R_{n_0}(\alpha_n)} \right|^2 \mu'(t)\, dm + \int_{\mathbb{T}} \left| \frac{\theta_{n-n_0}^* R_{n_0}}{\theta_{n-n_0}^*(\alpha_n) R_{n_0}(\alpha_n)} \right|^2 d\mu_s.$$

On the one hand, by properties (*ii*) and (*iii*) of Lemma 5.10, we get

$$\int_{\mathbb{T}} \left| \frac{\theta_{n-n_0}^* R_{n_0}}{\theta_{n-n_0}^*(\alpha_n) R_{n_0}(\alpha_n)} \right|^2 d\mu_s \leq \frac{\varepsilon^2 C_0^2}{(1-\varepsilon)^2} (\kappa'_{n-n_0}[\mu_{ac}])^{-2}. \tag{5.21}$$

On the other hand, by properties (*i*) and (*ii*) of the same lemma,

$$\begin{aligned} \int_{\mathbb{T}} \left| \frac{\theta_{n-n_0}^* R_{n_0}}{\theta_{n-n_0}^*(\alpha_n) R_{n_0}(\alpha_n)} \right|^2 \mu'(t)\, dm \quad &\leq \quad \frac{(1+\varepsilon)^2}{(1-\varepsilon)^2} \int_{\mathbb{T}\setminus\mathcal{W}_1} \left| \frac{\theta_{n-n_0}^*}{\theta_{n-n_0}^*(\alpha_n)} \right|^2 \mu'(t)\, dm \\ &\quad + \quad \frac{(2+\varepsilon)^2}{(1-\varepsilon)^2} \int_{\mathcal{W}_1} \left| \frac{\theta_{n-n_0}^*}{\theta_{n-n_0}^*(\alpha_n)} \right|^2 \mu'(t)\, dm. \end{aligned}$$

Expanding $(2+\varepsilon)^2$ and collecting terms, while using (5.20) and remembering that $\|\theta^*_{n-n_0}\|_{\mu_{ac}} = 1$, we obtain

$$\begin{aligned}\int_{\mathbb{T}} \left|\frac{\theta^*_{n-n_0} R_{n_0}}{\theta^*_{n-n_0}(\alpha_n) R_{n_0}(\alpha_n)}\right|^2 \mu'(t)\, dm &\leq \frac{(1+\varepsilon)^2}{(1-\varepsilon)^2}\left(\kappa'_{n-n_0}[\mu_{ac}]\right)^{-2} \\ &+ \frac{(3+2\varepsilon)\varepsilon}{(1-\varepsilon)^2}\left(\kappa'_{n-n_0}[\mu_{ac}]\right)^{-2}.\end{aligned} \tag{5.22}$$

Since ε can be made arbitrarily small, we gather from (5.21) and (5.22) that to each $\varepsilon' > 0$ there is n_0 such that

$$\kappa_n^{-2}[\mu] \leq (1+\varepsilon')\left(\kappa'_{n-n_0}[\mu_{ac}]\right)^{-2}$$

as soon as $n > n_0$. But from Proposition 5.9 we know that

$$\liminf_n (1-|\alpha_n|^2)|S(\alpha_n)|^2 \left(\kappa'_{n-n_0}[\mu_{ac}]\right)^2 \geq 1,$$

so we obtain (5.19) since ε' is arbitrarily small.

Next, we remove the assumption that $\mu' \geq \delta' > 0$ on $\mathcal{V}$ with $\mu' \in W^{1-1/q,q}$ on its components, but we suppose that $\mu' < C < \infty$ on $\mathcal{V}$. Fix a neighborhood $\mathcal{W}$ of $\operatorname{supp}\mu_s$ such that $\overline{\mathcal{W}} \subset \mathcal{V}$. To each $\eta > 0$, pick a neighborhood $\mathcal{V}_\eta \subset \mathcal{W}$ of $\operatorname{supp}\mu_s$ satisfying $|\mathcal{V}_\eta| < \eta$. Put $d\mu_\eta = \mu'_\eta dm + d\mu_s$, where $\mu'_\eta(t) = \mu'(t)$ for $t \notin \mathcal{V}_\eta$, $\mu'_\eta = C$ on $\mathcal{V}_\eta$. Being a positive constant on $\mathcal{V}_\eta$, μ'_η certainly meets the assumptions of the preceding part of the proof, so (5.19) holds for μ_η. Clearly, from (5.11), $\kappa_n[\mu_\eta] \leq \kappa_n[\mu]$ because $\mu \leq \mu_\eta$ hence

$$\liminf_n (1-|\alpha_n|^2)|S[\mu_\eta](\alpha_n)|^2 \kappa_n^2[\mu] \geq 1. \tag{5.23}$$

Now, let $\mathcal{V}'$ be open in $\mathbb{C}$ and contain no α_k, with $\mathcal{V}' \cap \mathbb{T} = \mathcal{V}$. Since

$$S[\mu](z) = S[\mu_\eta](z)\, \exp\left(\int_{\mathcal{V}_\eta} \frac{t+z}{t-z} \log|\mu'/C|\, dm(t)\right),$$

we see by dominated convergence (remember $\log\mu' \in L^1(\mathbb{T})$) that $S[\mu]/S[\mu_\eta]$ converges uniformly to 1 in $\mathbb{D} \setminus \mathcal{V}'$ as $\eta \to 0$. Consequently (5.19) follows from (5.23).

We now address the case where μ' may be unbounded in the neighborhood $\mathcal{V}$ of $\operatorname{supp}\mu_s$ but $\mu' \geq \delta'' > 0$ there. Since $p > 2$, $S[\mu]$ is continuous on $\mathcal{O}$ by Lemma 5.7. Observe also that $F_{\mu_{ac}} = \mu' + i\check{\mu}'$ lies in $W^{1-1/p,p}(I)$ for each component I of $\mathcal{O}$ and that F_{μ_s} is smooth on I, hence $F_\mu = F_{\mu_{ac}} + F_{\mu_s}$ is continuous and bounded on some open neighborhood $\mathcal{N}$ of $Acc(\alpha_k) \cap \mathbb{T}$ with $\overline{\mathcal{N}} \subset \mathcal{O}$. Moreover $|F_\mu| > \delta > 0$ on $\mathcal{O}$, thus $S[\widetilde{\mu}]$ is continuous and positively bounded from below on $\mathcal{N}$ by (5.18). Similarly $|F_\mu| \geq \delta''$ a.e. on $\mathcal{V}$, thus $\widetilde{\mu}' = \operatorname{Re} 1/F_\mu$ is bounded there. In addition, $\operatorname{supp}\widetilde{\mu}_s \cap \mathcal{N} = \emptyset$ otherwise the $H^\infty(\mathbb{D})$-function e^{-1/F_μ} would have a singular inner factor which is not analytic across $\mathcal{N}$, and its modulus could not be continuous and nonzero on $\mathcal{N}$ whereas $|F_\mu| \geq \delta$ there [18, Ch. II, Theorem 6.2]. Therefore we can apply the previous case of the proof to $\widetilde{\mu}$ with $\mathcal{O}$ replaced by $\mathcal{N}$; indeed, we know that $S[\widetilde{\mu}]$ is continuous on $\mathcal{N}$ which is enough to proceed by the remark after Proposition 5.9. Thus, we obtain for the ORFs of the second kind $\psi_n[\mu]$:

$$\lim_n |\psi_n^*[\mu](\alpha_n)|^2\, |S[\widetilde{\mu}](\alpha_n)|^2\, (1-|\alpha_n|^2) = 1. \tag{5.24}$$

Recalling from Proposition 2.4 that $\psi^*[\mu](\alpha_n)/\phi^*[\mu](\alpha_n) = F_\mu(\alpha_n)$, we conclude in view of (5.24) and (5.18) that (5.13) again holds.

Finally, under the sole assumptions of the proposition, let $\mathcal{V}$ be an open neighborhood of $\operatorname{supp}\mu_s$ in $\mathbb{T}$ such that $\overline{\mathcal{V}}\cap\mathcal{O}=\emptyset$, and $\mathcal{V}'$ be open in $\mathbb{C}$ and contain no α_k, with $\mathcal{V}'\cap\mathbb{T}=\mathcal{V}$. Let $\mathcal{W}$ be another neighborhood of $\operatorname{supp}\mu_s$ with $\overline{\mathcal{W}}\subset\mathcal{V}$. Put $d\mu_\varepsilon = \mu'_\varepsilon dm + d\mu_s$ with $\mu'_\varepsilon = \mu'+\varepsilon$ on $\mathcal{W}$, and $\mu'_\varepsilon = \mu'$ on $\mathbb{T}\backslash\mathcal{W}$. By what we just proved

$$\liminf_n (1-|\alpha_n|^2)|S[\mu_\varepsilon](\alpha_n)|^2\kappa_n^2[\mu_\varepsilon]\geq 1,$$

and since $\kappa_n[\mu_\varepsilon]\leq\kappa_n[\mu]$ (because $\mu_\varepsilon\geq\mu$) while

$$S[\mu]/S[\mu_\varepsilon] = \exp\left(\int_{\mathcal{W}}\frac{t+z}{t-z}\log|\mu'/(\mu'+\varepsilon)|\,dm(t)\right),\quad z\in\mathbb{D},$$

converges uniformly to 1 in $\mathbb{D}\setminus\mathcal{V}'$ as $\varepsilon\to 0$ by the monotone convergence of $\mu'/(\mu'+\varepsilon)$ to 1 a.e. on $\mathcal{W}$, we conclude that (5.19) holds. □

In the course of the previous proof, we noticed that (5.24) is equivalent to (5.13). This is worth recording, taking into account that $\widetilde{\widetilde{\mu}}=\mu$:

Corollary 5.12. *Let $\mu\in$ (S). Then (5.13) holds for μ if, and only if it holds for $\widetilde{\mu}$, the Herglotz measure of $1/F_\mu$ (see (0.4)).*

There are Carathéodory functions, with continuous and strictly positive real part on $\mathbb{T}$, whose imaginary part is unbounded. One example is $2+\varphi$ where φ conformally maps $\mathbb{D}$ onto $\{z=x+iy;\ |x|<1/(1+y^2)\}$, $\varphi(0)=0$ and $\varphi'(0)>0$, whose imaginary part is unbounded at $\pm i$, see [18, Ch. III, Sect. 1]. If we put $d\mu'(t)=(2+\operatorname{Re}\varphi(t))dt$, then $2+\varphi(t)=F_\mu$ and and $\widetilde{\mu}'=\mu'/|F_\mu|^2$ is continuous but vanishes at $\pm i$. Letting (α_k) accumulate at $\pm i$, Theorem 5.8 will apply to μ and then Corollary 5.12 will provide us with an example where (5.13) holds although (0.12) fails.

Proof of Theorem 5.8. Let $\mathcal{O}$ be the neighborhood of $Acc(\alpha_k)\cap\mathbb{T}$ granted by (0.11)-(0.13). Shrinking $\mathcal{O}$ if necessary, we may assume that μ' is continuous with $\mu'\geq\delta>0$ on a neighborhood of $\overline{\mathcal{O}}$ in $\mathbb{T}$. Pick $\varepsilon>0$ and $0<r<1$ so that the Poisson integral $h_r(z)=P_{rz}*\mu'$ satisfies $|h_r-\mu'|<\varepsilon$ on $\overline{\mathcal{O}}$. Let μ_ε have singular part μ_s and absolutely continuous part $\mu'_\varepsilon\,dm$ where $\mu'_\varepsilon(t)=\mu'(t)$ for $t\notin\mathcal{O}$ and $\mu'_\varepsilon(t)=h_r(t)+\varepsilon$ for $t\in\mathcal{O}$. Then $\mu'\leq\mu'_\varepsilon\leq\mu'+2\varepsilon$ on $\mathbb{T}$ and μ'_ε is smooth on $\mathcal{O}$. By Proposition 5.11, we have

$$\lim_n\kappa_n^2[\mu_\varepsilon]\,|S[\mu_\varepsilon](\alpha_n)|^2\,(1-|\alpha_n|^2)=1. \tag{5.25}$$

Since $\kappa_n[\mu]\geq\kappa_n[\mu_\varepsilon]$ because $\mu\leq\mu_\varepsilon$, we deduce from (5.25) that

$$\begin{aligned}&\liminf_n(1-|\alpha_n|^2)|S(\alpha_n)|^2\kappa_n^2[\mu]\\ &\geq\liminf_n\frac{|S(\alpha_n)|^2}{|S[\mu_\varepsilon](\alpha_n)|^2}\,(1-|\alpha_n|^2)|S[\mu_\varepsilon](\alpha_n)|^2\kappa_n^2[\mu_\varepsilon]\\ &=\liminf_n\frac{|S(\alpha_n)|^2}{|S[\mu_\varepsilon](\alpha_n)|^2}.\end{aligned}\tag{5.26}$$

Recalling the inequalities on μ',μ'_ε given above, we get

$$\frac{|S(z)|}{|S[\mu_\varepsilon](z)|}=\exp\bigl(P_z*\log(\mu'/\mu'_\varepsilon)\bigr)\geq 1-2\varepsilon/\delta \tag{5.27}$$

for $z \in \mathbb{D}$, and letting $\varepsilon \to 0$ we obtain (5.19) from (5.26), (5.27). With Corollary 5.3, we finish the proof. At last, assume that (α_k) accumulates nontangentially on $Acc(\alpha_k) \cap \mathbb{T}$ and that (0.11), (0.12) get replaced by (0.15). Then we can find a sequence of continuous functions $\varphi_j > 0$ decreasing pointwise to μ' on $\mathcal{O}$. Letting $d\mu_j = d\mu_s + \mu_j' dm$ where $\mu_j' = \varphi_j$ on $\mathcal{O}$ and $\mu_j' = \mu'$ on $\mathbb{T} \setminus \mathcal{O}$, we get from the first part of the proof that (5.13) holds for μ_j. Since $\mu_j \geq \mu$, we deduce as in (5.26) that for each j

$$\liminf_n (1 - |\alpha_n|^2)|S(\alpha_n)|^2 \kappa_n^2[\mu] \geq \liminf_n \frac{|S(\alpha_n)|^2}{|S[\mu_j](\alpha_n)|^2}. \tag{5.28}$$

Without loss of generality, we may assume that (α_n) converges to $\alpha \in \overline{\mathbb{D}}$. If $\alpha \in \mathbb{D}$, the conclusion follows from the fact that $S[\mu_j](\alpha) \to S(\alpha)$ by the monotone convergence of μ_j' to μ'. If $\alpha \in \mathbb{T}$, then by Fatou's theorem

$$\lim_n \frac{|S(\alpha_n)|^2}{|S[\mu_j](\alpha_n)|^2} = \frac{\mu'(\alpha)}{\mu_j'(\alpha)},$$

which can be made arbitrarily close to 1 since $\lim_j \mu_j'(\alpha) = \mu'(\alpha) > 0$. $\square$

Corollary 5.13. *Let* (0.3), (0.11)-(0.13) *be satisfied and* $\mu \in$ (S). *Then*

$$\lim_n \left\| S\phi_n^*(z) - \beta_n \frac{\sqrt{1 - |\alpha_n|^2}}{1 - \overline{\alpha}_n z} \right\| = 0, \tag{5.29}$$

where the unimodular factors β_n *are defined in Theorem 3' of Section 0.2. Moreover, for any sequence* $(z_n) \subset \mathbb{D}$*, it holds that*

$$\lim_n \left\{ \phi_n^*(z_n) S(z_n) \sqrt{1 - |z_n|^2} - \beta_n \frac{\sqrt{1 - |\alpha_n|^2}\sqrt{1 - |z_n|^2}}{1 - \overline{\alpha}_n z_n} \right\} = 0. \tag{5.30}$$

If (α_k) *accumulates nontangentially on* $Acc(\alpha_k) \cap \mathbb{T}$*, then it is enough to assume instead of* (0.11)-(0.12) *that* (0.15) *holds.*

Proof. Estimating the integral, we get

$$\begin{aligned}
\left\| S\phi_n^*(z) - \beta_n \frac{\sqrt{1 - |\alpha_n|^2}}{1 - \overline{\alpha}_n z} \right\|^2 &= \int_{\mathbb{T}} \left| S\phi_n^*(z) - \beta_n \frac{\sqrt{1 - |\alpha_n|^2}}{1 - \overline{\alpha}_n z} \right|^2 dm(z) \\
\leq \|\phi_n\|_\mu^2 &- 2\mathrm{Re}\left(\frac{\overline{\beta}_n}{2i\pi} \int_{\mathbb{T}} S\phi_n^*(z) \frac{\sqrt{1 - |\alpha_n|^2}}{z - \alpha_n} dz \right) + 1 \\
&= 2(1 - \sqrt{1 - |\alpha_n|^2}\, |S(\alpha_n)|\, |\phi_n^*(\alpha_n)|),
\end{aligned}$$

and Theorem 5.8 yields (5.29). Next, let us set

$$k_{z_n} = \frac{\sqrt{1 - |z_n|^2}}{1 - z\overline{z}_n}, \qquad G_n(z) = S\phi_n^*(z) - \beta_n \frac{\sqrt{1 - |\alpha_n|^2}}{1 - \overline{\alpha}_n z}.$$

Since $\|k_{z_n}\| = 1$, the relation just proven and the Schwarz inequality yield $\lim_n (G_n, k_{z_n}) = 0$. Expanding the scalar product gives us (5.30). $\square$

Under the assumptions of the theorem, its conclusions also hold for $\tilde{\mu}$ by Corollary 5.12.

Corollary 5.14. *Let* (0.3) *hold and* $\mu \in$ (S) *meet* (0.11)-(0.13)*. If I is an open arc on $\mathbb{T}$, with $\overline{I} \cap \operatorname{supp} \mu_s = \emptyset$, such that $\mu' \in W^{1-1/p,p}(I)$ for some $p > 4$, then $F_\mu \phi_n^* - \psi_n^*$ converges to zero locally uniformly on I.*

Proof. From Corollary 5.13, limit (5.29) holds as well as its analogue for $\widetilde{\mu}$:

$$\lim_n \left\| S[\widetilde{\mu}]\psi_n^*(z) - \beta_n[\widetilde{\mu}] \frac{\sqrt{1-|\alpha_n|^2}}{1-\overline{\alpha}_n z} \right\| = 0. \tag{5.31}$$

Moreover, it follows immediately from (5.18) and Proposition 2.4 that $\beta_n[\mu] = \beta_n[\widetilde{\mu}]$. Substracting (5.31) from (5.29) and using (5.18) now gives us

$$\lim_n \|S[\widetilde{\mu}](F_\mu \phi_n^* - \psi_n^*)\| = \lim_n \|S[\mu]\phi_n^* - S[\widetilde{\mu}]\psi_n^*\| = 0.$$

In particular, we get that from any subsequence of $g_n := F_\mu \phi_n^* - \psi_n^*$ one can extract a subsequence that converges pointwise a.e. to zero on $\mathbb{T}$. But since g_n is equicontinuous on compact subsets of I by Proposition 5.6, *ii*), we deduce from Ascoli's theorem that g_n converges locally uniformly to zero on I. □

Acknowledgments. The second author is grateful to members of the APICS team from INRIA Sophia-Antipolis for numerous invitations and warm hospitality.

References

[1] Adams, R., Fournier J. Sobolev spaces. Pure and Applied Maths., vol. 140, Academic Press, 2003.
[2] Akhiezer, N. Theory of approximation. Dover Publications Inc., New York, 1992.
[3] Anderson, B., Vongpanitlerd, S. Network Analysis and Synthesis - A modern system theory approach. Prentice Hall, 1973.
[4] Aptekarev, A. Sharp constant for rational approximation of analytic functions. Math. Sb. 193 (2002), no. 1-2, 1–72.
[5] Aptekarev A., Van Assche, W. Scalar and Matrix Riemann-Hilbert approach to the strong asymptotics of Padé approximants and complex orthogonal polynomials with varying weight. J. Approx. Theory 129 (2004), 129–166.
[6] Astala K., Iwaniec T., Martin G. Elliptic partial differential equations and quasiconformal mappings in the plane. Math. Series 48, Princeton Univ. Press, 2009.
[7] Baratchart, L., Küstner, R., Totik, V. Zero distribution via orthogonality. Ann. Inst. Fourier 55 (2005), 1455–1499.
[8] Baratchart, L., Yattselev, M. Convergent interpolation to Cauchy integrals over analytic arcs. to appear in Found. Constr. Math.
[9] Begehr, H. Complex analytic methods for partial differential equations. World Scientific, 1994.
[10] Bultheel, A., González-Vera, P., Hendriksen, E., Njåstad, O. Orthogonal rational functions. Cambridge University Press, Cambridge, 1999.
[11] Bultheel, A., González-Vera, P., Hendriksen, E., Njåstad, O. Orthogonal rational functions on the unit circle: from the scalar to the matrix case. Lecture Notes in Math., vol. 1883 (2006), 187–228.
[12] Campanato, S. Elliptic systems in divergence form, Interior regularity, Quaderni, Scuola Normale Superiore Pisa, 1980.
[13] Dzrbasjan, M. Orthogonal systems of rational functions on the unit circle with given set of poles. Dokl. Akad. Nauk. SSSR 147 (1962), 1278–1281 (Russian); English transl. in Soviet Mat. Dokl., (1962) no. 3, 1794–1798.
[14] Denisov, S., Kupin, S. Orthogonal polynomials and a generalized Szegő condition. C. R. Math. Acad. Sci. Paris 339 (2004), no. 4, 241–244.

[15] Denisov, S., Kupin, S. Asymptotics of the orthogonal polynomials for the Szegő class with a polynomial weight. J. Approx. Theory 139 (2006), no. 1-2, 8–28.
[16] Duren, P. Theory of H^p spaces. Academic Press, New York, 1970.
[17] Faurre, P., Clerget, M., Germain, F. Operateurs rationels positifs. Dunod, 1979.
[18] Garnett, J. Bounded analytic functions. Springer, New York, 2007.
[19] Geronimus, J. On polynomials orthogonal on the circle, on trigonometric moment-problem and on allied Carathéodory-Schur functions. Rec. Math. [Mat. Sbornik] N. S. 15 (1944), no. 57, 99–130.
[20] Geronimus, J. Orthogonal polynomials. Consultants Bureau, New York, 1961.
[21] Grisvard, P. Elliptic problems in nonsmooth domains. *Monographs and Studies in Mathematics* 24, Pitman, 1985.
[22] Khrushchev, S. Schurs algorithm, orthogonal polynomials, and convergence of Wall's continued fractions in $L^2(T)$. J. Approx. Theory 108 (2001), no. 2, 161–248.
[23] Khrushchev, S. Classification theorems for general orthogonal polynomials on the unit circle. J. Approx. Theory 116 (2002), no. 2, 268–342.
[24] Killip, R., Simon, B. Sum rules for Jacobi matrices and their applications to spectral theory. Ann. of Math. (2) 158 (2003), no. 1, 253–321.
[25] Koosis, P. Introduction to H^p spaces. Cambridge University Press, Cambridge, 1998.
[26] Krein, M. On a generalization of some investigations of G. Szegő, V. Smirnoff and A. Kolmogoroff. C. R. (Doklady) Acad. Sci. URSS (N.S.) 46 (1945). 91–94.
[27] Kuijlaars, A., McLaughlin, K., Van Assche, W., Vanlessen, M. The Riemann-Hilbert approach to strong asymptotics for orthogonal polynomials on $[-1, 1]$. Adv. Math. 188 (2004), no. 2, 337–398.
[28] Langer, H., Lasarow, A. Solution of a multiple Nevanlinna-Pick problem via orthogonal rational functions. J. Math. Anal. Appl. 293 (2004), no. 2, 605–632.
[29] Lubinsky, D. Universality limits in the bulk for arbitrary measures on a compact set, to appear in J. d'Analyse Math.
[30] Lunot, V. Rational approximation techniques and frequency design: a Zolotarev problem and the Schur algorithm, Ph.D. Thesis, University of Provence, 2008.
[31] McLaughlin, K., Miller, P. The $\bar{\partial}$ steepest descent method for orthogonal polynomials on the real line with varying weight, submitted.
[32] Martínez-Finkelstein, A., McLaughlin, K., Saff, E. Asymptotics of orthogonal polynomials with respect to an analytic weight with algebraic singularities on the circle. Int. Math. Res. Notices ID 91426 (2006), 1-43.
[33] Matthaei, Y. Microwave filters, impedance matching networks and coupling structures. New York, Mc Graw Hill, 1965.
[34] Miña-Dìaz, E. An expansion for polynomials orthogonal over an analytic Jordan curve. Comm. Math. Phys. 285 (2009), no. 3, 1109–1128.
[35] Nevanlinna, R. Über beschränkte Funktionen, die in gegebenen Punkten vorgeschriebene Werte annehemen. Ann. Acad. Sci. Fenn. 13 (1919) no. 1.
[36] Nikishin, E., Sorokin, V. Rational approximations and orthogonality. Translations of AMS, vol. 92, Providence, RI, 1991.
[37] Njastad, O., Velàzquez, L. Wall rational functions and Khrushchev's formula for orthogonal rational functions. Constr. Approx. 30 (2009), no. 2, 277-297.
[38] Pan, K. On the convergence of rational functions orthogonal on the unit circle. J. Comput. Appl. Math. (1996), no. 76, 315–324.
[39] Peherstorfer, F., Yuditskii, P. Asymptotics of orthonormal polynomials in the presence of a denumerable set of mass points. Proc. Amer. Math. Soc. 129 (2001), no. 11, 3213–3220.
[40] Rakhmanov, E. Asymptotic properties of polynomials orthogonal on the unit circle with weights not satisfying the Szegő condition (Russian). Mat. Sb. 130 (172) (1986), no. 2, pp. 151-169, 284.
[41] Remling, C. The absolutely continuous spectrum of Jacobi matrices, submitted.
[42] Schur, I. Über potenzreihen, die im innern des einheitskreises beschränkt sind. J. Reine Angew. Math. 147 (1917) 205232. English translation in: I. Schur methods in operator theory and signal processing (Operator Theory: Adv. and Appl. 18 (1986), Birkhäuser Verlag).

[43] Simon, B. Orthogonal polynomials on the unit circle, I, II. AMS Colloquium Publications, vol. 54, Providence, RI, 2005.
[44] Szegő, G. Orthogonal polynomials. AMS, Providence, RI, 1975.
[45] Wall, H. Continued fractions and bounded analytic functions. Bull. Amer. Math. Soc. 50 (1944). 110–119.
[46] Wall, H. Analytic theory of continued fractions. Van Nostrand, New York, 1948.
[47] Ziemer, W. P. Weakly differentiable functions. G.T.M. 120, Springer, 1989.

2004 route des Lucioles - BP 93 FR-06902 Sophia Antipolis Cedex, France
E-mail address: Laurent.Baratchart@sophia.inria.fr

IMB, Université Bordeaux 1, 351 cours de la Libération, 33405 Talence Cedex France
E-mail address: skupin@math.u-bordeaux1.fr

2004 route des Lucioles - BP 93 FR-06902 Sophia Antipolis Cedex, France
E-mail address: vlunot@gmail.com

2004 route des Lucioles - BP 93 FR-06902 Sophia Antipolis Cedex, France
E-mail address: Martine.Olivi@sophia.inria.fr